A RADIO POLARIMETER SPECTROMETER

by

Peter John Hall, B.Eng., B.Sc.(Hons.)

Submitted in fulfilment of the
requirements for the degree of
Doctor of Philosophy

UNIVERSITY OF TASMANIA
HOBART
MARCH, 1985

to graduate
1985.

This thesis contains no material which has been accepted for the award of any other degree or diploma in any university. To the best of my knowledge and belief, the thesis contains no copy or paraphrase of material previously published or written by another person except where due reference is made in the text.

P. J. Hall

TABLE OF CONTENTS

SUMMARY

For several years it has been apparent that a requirement in radioastronomy is a device capable of providing high time resolution polarization spectra of a wide bandwidth. This thesis describes such an instrument. The new polarimeter spectrometer is based on two surface acoustic wave (SAW) spectrum analysers operating in a coherent (phase preserving) mode. The method of spectrum analysis is a high sensitivity one, the analysers being more akin to filterbank types than to swept-frequency instruments.

A wideband analog multiplier has been developed, permitting the spectrum analyser outputs to be correlated directly to yield cross-spectral or polarization data. As well, the usual problem with SAW analysers - the high output data rate - has been overcome with the development of a fast digital spectrum integrator. This device allows the SAW intermediate-frequency processor to be interfaced to a general purpose computer, resulting in a real-time signal processing system with great versatility.

The prototype SAW processor produces polarization spectra over a bandwidth of about 30 MHz with frequency and time resolutions of 650 kHz and 2.4 μs. When interfaced to a typical computer, Stokes polarization spectra with a time resolution of about 150 μs are available. After allowing for non-continuous operation of the prototype spectrum analysers, the radiometer efficiency of the complete processing system exceeds 70%. Most techniques used in the development of the first polarimeter spectrometer are applicable to larger scale instruments.

Although primarily the report on a project in applied signal processing, this thesis contains three other main sections. Firstly, an extensive and up-to-date review of SAW spectrum analysis and associated techniques is included. Secondly, the results of a series of numerical simulations are given. The simulations relate to the performance of post-detection digital integrators when quantizing envelope-detected, inverse bandwidth noise.

In the third major section of the thesis, data from initial scientific observations conducted using the new instrumentation are

presented. The SAW and digital processors, together with other items of hardware constructed in the course of the research project, were used in conjunction with the 14 m steerable radio telescope at the Llanherne (Hobart) observatory. As well as confirming laboratory radiometer efficiency measurements, the first astronomical data from the new system have been used to demonstrate changes in the rotation and dispersion measures of the Vela pulsar, PSR 0833-45.

The rotation measure change is in general agreement with a trend deduced by earlier workers. The change in dispersion measure is only the second confirmed for any pulsar, the first changes being reported more than 10 years ago for the Crab nebula pulsar. The present variation is over 100 times the order of the Crab pulsar changes and amounts to more than 1% of the Vela pulsar dispersion measure reported by investigators in 1970. The new data support the hypothesis of relative motion between PSR 0833-45 and a magnetized filament in the Vela supernova remnant.

ACKNOWLEDGEMENTS

I thank my supervisor Dr. P.A. Hamilton for his guidance throughout this project. I am also indebted to Professor G.R.A. Ellis for his advice and continual encouragement and to Dr. P.M. McCulloch for useful discussions on many matters.

It is a pleasure to acknowledge the assistance of Mr. M.E. Costa in the writing of software for several tasks undertaken during the final stages of the project. In particular, his help in the numerical simulation experiments and the Llanherne observations expedited greatly the last stages of the work.

I am grateful to the technical and workshop staff in the Physics Department for their assistance at various times. My special thanks go to Mr. P.E. Button for constructing the main computer interface and to Mr. G.A. Gowland for helpful discussions.

The comments of many people outside the University of Tasmania have been invaluable during the project. Dr. J.G. Ables, Professor T.W. Cole, Mr. B.F.C. Cooper, Dr. R.H. Frater, Professor D.G. Lampard and Dr. A.G. Lyne have all given freely of their time to discuss aspects of the work. Similarly, discussions with Dr. J.W. Arthur, Dr. P.M. Grant and Mr. G.L. Moule on matters related to SAW spectrometer design and performance have been fruitful. I thank all of these people.

The support of my parents and colleagues over the past four years has been greatly appreciated and I record here my sincerest appreciation.

Finally, I thank Mrs. Betty Golding for her expert and speedy typing of this thesis.

During 1981 and 1982 I received a Commonwealth Postgraduate Research Award. For the past two years I have held a Radio Research Board Postgraduate Scholarship. A visit to the U.K. during the course of the project was made possible by my receipt of the 1982 University of Tasmania H.M. Bamford Travel Scholarship.

CHAPTER 1 INTRODUCTION

1.1 INTRODUCTION

Radioastronomy is the study of the radiation of particles or groups of particles in extra-terrestial environments. When the motion of particles is ordered by an external constraint such as a magnetic field their natural radiation becomes polarized. The study of polarization is therefore invaluable in reaching conclusions about ordering influences within radio sources. Furthermore, the polarization state of radio waves is modified when the radiation passes through plasma propagation media, so polarization studies are often useful in establishing properties of cosmic plasmas.

For some years it has been apparent that a major requirement in the study of non-stationary (time varying) cosmic radiation is an instrument able to provide high time resolution polarization spectra over a wide bandwidth. The main purpose of this research project has been to examine, in a practical way, the feasibility of constructing polarimeter spectrometers based on a relatively new method of real-time Fourier transformation or coherent spectrum analysis. The technique is implemented using surface acoustic wave (SAW) dispersive filters.

SAW spectrum analysis has not been used before in radioastronomy but the basic techniques have been applied by the military. The signal processing methods detailed in this thesis extend considerably the elementary approaches to spectrum analysis and post-transformation processing employed by previous users.

Two new items of equipment are described in the thesis. The basis of the signal processing system is a SAW polarimeter spectrometer or coherent IF processor which provides polarization spectra for a 30 MHz band with time and frequency resolutions of 2.4 μs and 650 kHz. An essential part of the instrument is a wideband, four-quadrant analog multiplier which was developed for the application. The multiplier has input and output bandwidths in excess of 50 MHz and allows direct correlation of the outputs of two SAW spectrum analysers.

The second major item of equipment is a high speed digital spectrum integrator designed to slow the very high output data rate

of SAW spectrum analysers to one compatible with laboratory computers. The new integrator can sample and process data in real-time with a clock frequency of up to 30 MHz. The combination of a SAW power spectrum analyser and integrator will be referred to as a SAW radiometer, while the term "SAW coherent radiometer" has been adopted to describe the IF processor and its four associated spectrum integrators.

In order to complete a scientifically useful system, several other pieces of apparatus have been constructed and these are described briefly. A pulsar timing unit (PTU) permits the radiometer to operate synchronously with pulsar signals and a wideband, dual-channel receiver is an essential part of the observing system. A direct memory access (DMA) computer interface and a high speed data link enable the radiometer to communicate with common mini-computers.

As well as descriptions of equipment this thesis contains three other main sections. Firstly, an extensive review of SAW spectrum analysis and associated techniques has been compiled. Secondly, the results of a series of numerical simulation experiments are reported. Simulation techniques have been used to quantify the dynamic range and signal-to-noise ratio performance of digital integrators processing envelope-detected, inverse bandwidth (un-smoothed) noise. This topic arises from a consideration of the requirements for quantizing the output of a SAW spectrum analyser.

In the remaining major section of the thesis the first scientific data obtained using the new instrumentation are presented. These data reveal changes in the dispersion and rotation measures of the Vela pulsar, PSR 0833-45. The magnitude of the rotation measure change is in general agreement with a trend deduced by earlier investigators. The change in dispersion measure is the second reported for a pulsar, several small changes having been detected more than ten years ago in the Crab pulsar dispersion measure. The latest change is more than 1% of the dispersion measure recorded for the Vela pulsar in 1970 and is over 100 times the magnitude of the Crab pulsar changes. The new results are consistent with the hypothesis of relative motion between PSR 0833-45 and a magnetized region of excess electron density lying in the line of sight to the pulsar. The values deduced for the scale size and internal magnetic field of the region suggest that it is probably part of a magnetized filament in the Vela supernova remnant.

The thesis conclusion is that the coherent radiometer is a workable scientific instrument having a noise efficiency of more than 70% after allowance is made for non-continuous operation of the prototype spectrum analysers. Several major practical problems associated with the application of SAW signal processors have been overcome and this, together with the success of the initial observations, justifies the claim that SAW processing has considerable potential in radioastronomy.

High speed Fourier transformation and transform domain correlation is used extensively in many areas of science and engineering. Descriptions in this thesis relate directly to coherent spectrum analysis and polarimetry but sufficient information is given to allow assessment of the role of SAW technology in other areas. For example, the prototype instrumentation could provide high time resolution correlation outputs from a broadband radio interferometer. Similarly, the system could function as a fast surveillance receiver able to provide both signal time and angle of arrival information.

Two appendices to the thesis cover the basic theory of SAW devices as well as the design of SAW dispersive filters. Commercial SAW filters were used in the prototype processor but it is feasible for research institutions with access to micro-electronic fabrication facilities to make SAW devices.

The remainder of chapter one contains review material relating to methods of specifying and measuring partially polarized electromagnetic waves, the effects of propagation media on cosmic radio signals, and the properties of non-stationary radiation from Jupiter, the sun and pulsars. A pictorial signal-space representation of cosmic radio signals is also discussed briefly.

1.2 WAVE POLARIZATION

Cosmic radio waves are broadband signals. Electrical fluctuations induced by the waves in antennas are amplified and limited ultimately to some observing bandwidth Δf, allowing the then quasi-monochromatic noise disturbance to be written in the narrowband form

$$y(t) = A(t) \exp\{j[\omega t + \phi(t)]\}. \qquad 1.2.1$$

The envelope $A(t)$ and phase $\phi(t)$ both vary on a time-scale $t \sim 1/\Delta f$ which is slow compared with the harmonic centre frequency variation. The signal is purely real, the complex form used in equation 1.2.1 being an analytic signal representation (Ko, 1962) in which the real part is equal to the real function and the imaginary part is in quadrature with the real part.

Suppose quasi-monochromatic radiation travels in the +z direction in an xyz right-handed co-ordinate system. In some fixed z-plane, x and y components of the wave are:

$$E_x(t) = E_1(t) \exp\{j[\omega t - \delta_1(t)]\}$$
$$E_y(t) = E_2(t) \exp\{j[\omega t - \delta_2(t)]\} \; . \qquad 1.2.2$$

Depending on the degree of coherence between the orthogonal components, the wave is said to be unpolarized (randomly polarized), partially polarized or fully polarized. The partially polarized case is the general one and has been discussed in detail by Cohen (1958), Born and Wolf (1970), Kraus (1966) and Longhurst (1973).

Two equivalent methods are used to describe partially polarized radiation. In the first method the component intensities are given and the correlation (coherence) between components is specified. The alternative method involves specification of the relative intensities of the polarized and unpolarized fractions as well as the polarization state of the polarized fraction. The first approach is now considered.

Wolf (1959) has defined a coherency matrix,

$$C = \begin{bmatrix} C_{xx} & C_{xy} \\ C_{yx} & C_{yy} \end{bmatrix}$$

which, for the expressions in equations 1.2.2 simplifies to:

$$C = \begin{bmatrix} \langle E_1^2 \rangle & \langle E_1 E_2 e^{j\delta} \rangle \\ \langle E_1 E_2 e^{-j\delta} \rangle & \langle E_2^2 \rangle \end{bmatrix}$$

where $\delta = \delta_1 - \delta_2$, $\langle\ \rangle$ indicates a time average and the product terms are cross-correlation functions of the quasi-monochromatic variations. Note that $C_{xy} = C_{yx}^*$, so there are four independent scalar parameters: C_{xx}, C_{yy}, and the modulus and argument of either C_{yx} or C_{xy}. The component intensities are C_{xx} and C_{yy} and the total intensity is $C_{xx} + C_{yy}$. Wolf also defines a complex correlation factor:

$$\rho_{xy} = |\rho_{xy}| \exp(j\beta_{xy}) = \frac{C_{xy}}{\sqrt{C_{xx}}\sqrt{C_{yy}}} \quad . \qquad 1.2.3$$

Using the Schwartz inequality, it can be shown that $|\rho_{xy}| \leq 1$. In fact, $|\rho_{xy}|$ is a measure of the degree of coherence (polarization) and β_{xy} (δ for the signals defined in equation 1.2.2) is the effective phase difference between components.

As an alternative to the coherency matrix, partially polarized radiation can be specified by a four-element vector having real components. Elements of the vector are known as Stokes parameters and are defined as follows:

$$\begin{aligned} I &= \langle E_1^2 \rangle + \langle E_2^2 \rangle = [C_{xx} + C_{yy}] \\ Q &= \langle E_1^2 \rangle - \langle E_2^2 \rangle = [C_{xx} - C_{yy}] \\ U &= 2\langle E_1 E_2 \cos\delta \rangle = [C_{xy} + C_{yx}] \\ V &= 2\langle E_1 E_2 \sin\delta \rangle = -j[C_{xy} - C_{yx}] \quad . \end{aligned} \qquad 1.2.4$$

Returning to the alternative description of partially polarized radiation, the Stokes parameters can be written in terms of polarized and unpolarized fractions. If A_1 and A_2 are amplitudes of the polarized components and A_u is the amplitude of the unpolarized fraction, the parameters are:

$$I = A_1^2 + A_2^2 + A_u^2$$

$$Q = A_1^2 - A_2^2$$

$$U = 2\, A_1 A_2 \cos\delta \qquad 1.2.5$$

$$V = 2\, A_1 A_2 \sin\delta \quad .$$

In both equations 1.2.4 and 1.2.5, I is the total intensity and Q is the difference between component intensities. The incoherent fraction of the radiation contributes nothing to the U or V time averages in equations 1.2.4, leaving the values in equations 1.2.5. The parameters U and V are orthogonal projections of the zero-lag cross-correlation vector, and I and Q can be regarded as the sum and difference of two zero-lag autocorrelation products.

For completely polarized radiation

$$I^2 = Q^2 + U^2 + V^2 \ , \qquad 1.2.6$$

while for partially polarized signals

$$I^2 \geq Q^2 + U^2 + V^2 \ .$$

The degree of polarization is defined as

$$m = \frac{\text{intensity of polarized fraction}}{\text{total intensity}} = \frac{\sqrt{Q^2+U^2+V^2}}{I} \ .$$

Equation 1.2.6 defines a sphere (the Poincaré sphere) of radius I if Q, U, and V are taken as lying along orthogonal axes in a Cartesian co-ordinate system. In this representation (figure 1-1(a)) every possible polarization state has a corresponding point on the sphere. Figure 1-1(b) shows in detail the polarization ellipse (the two-dimensional representation of the motion of the wave electric field vector) at one point on the sphere. Born and Wolf (1970) show

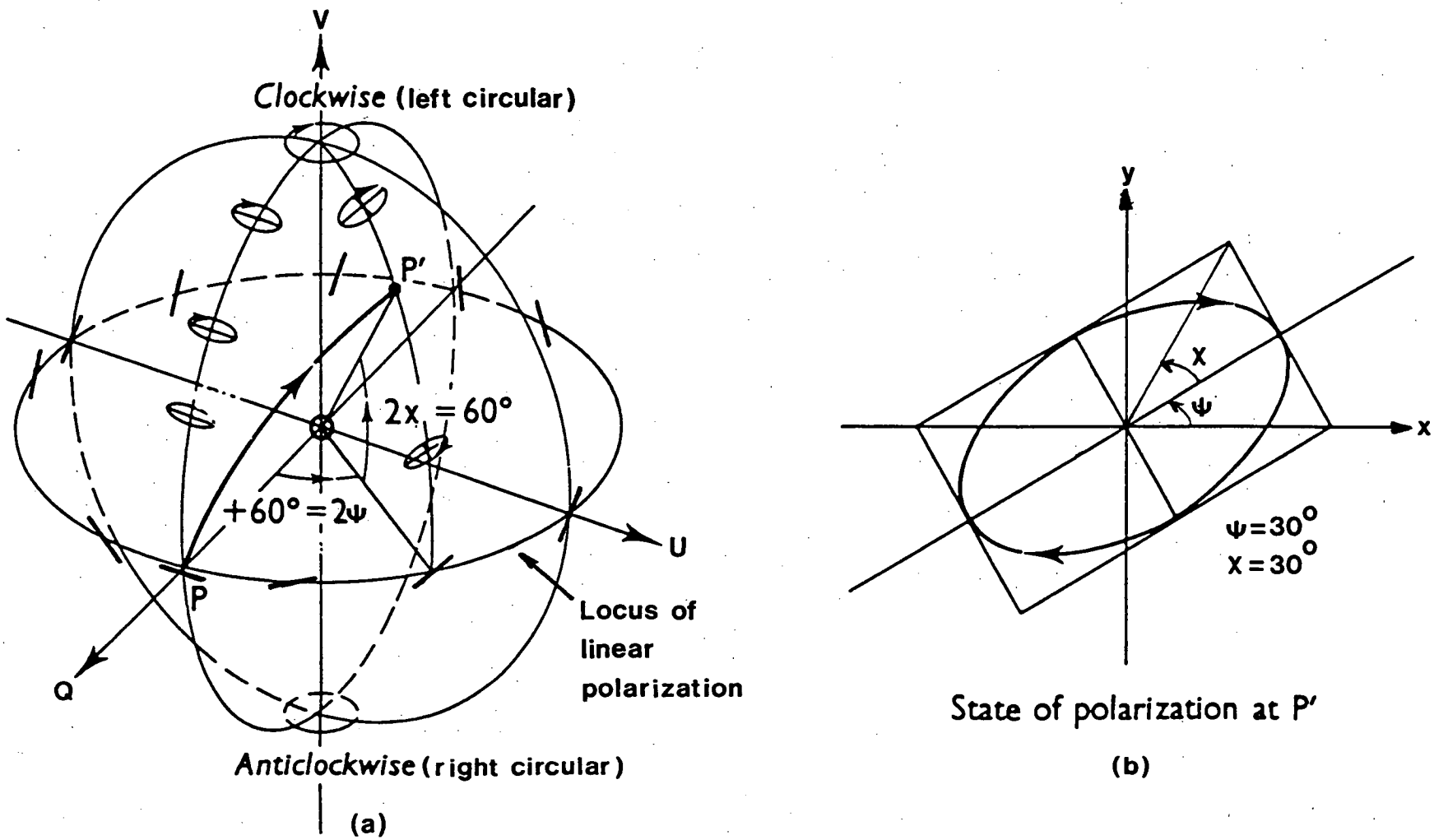


Fig.1-1: Poincaré sphere representation of polarized electromagnetic radiation. In (a) moving from point P to P' on the sphere corresponds to a change from linear to elliptical polarization. The polarization ellipse at point P' is shown in (b). (Adapted from Longhurst, 1973).

that the angles ψ and χ are related to the Stokes parameters by:

$$Q = I \cos 2\chi \cos 2\psi$$
$$U = I \cos 2\chi \sin 2\psi$$
$$V = I \sin 2\chi \quad .$$

The angle $\psi = \frac{1}{2} \arctan (U/Q)$ is the position angle (PA) of the linearly polarized component and the linearly polarized intensity is

$$L = \sqrt{Q^2 + U^2} \quad .$$

The fractional linear polarization is L/I. The parameter V is the circularly polarized intensity, the fractional circular polarization being $|V|/I$. The Poincaré representation can be retained for the polarized fraction of partially polarized radiation if the intensity vector is taken to be the polarized intensity.

1.3 MEASUREMENT OF STOKES PARAMETERS

The Stokes parameters are a convenient means of specifying partially polarized radiation because they can be measured using conventional antenna and RF engineering techniques. In the basic measurement system, two antennas sensitive to opposite polarizations are required. Any wave can be regarded as the sum of two independent, elliptically polarized constituents having contra-rotating field vectors and perpendicular polarization ellipses. These two waves are said to be oppositely polarized. Perpendicular linear antennas and opposite-hand circular antennas are examples of oppositely sensitive antennas (Cohen, 1958).

If linear antennas are used in the parameter measurement, the operations prescribed in equations 1.2.4 can be performed using standard RF hardware (figure 1-2). In the resulting instrument, known as a narrowband correlation Stokes polarimeter, signals from perpendicular antennas are mixed to an IF band. Adding and subtracting the square-law detected (intensity) outputs gives I and Q. To obtain U and V, in-phase and quadrature correlation is performed using four-quadrant analog multipliers. The necessary time averaging is done by RC networks.

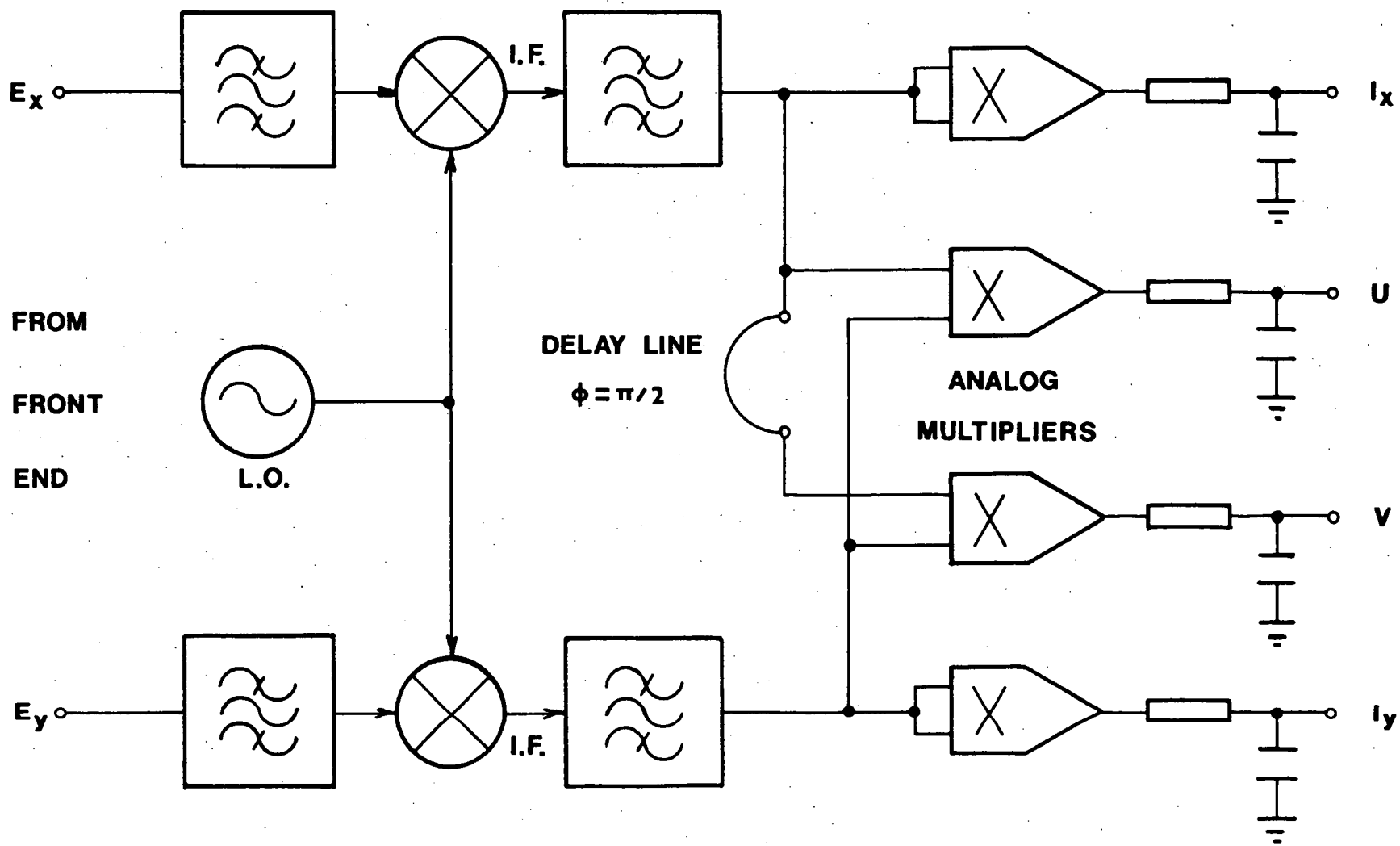


Fig.1-2: Narrowband correlation Stokes polarimeter.

If opposite-hand circular antennas are used, Cohen (1958) shows that equivalent expressions are:

$$\begin{aligned} I &= \langle E_3^2 \rangle + \langle E_4^2 \rangle \\ Q &= 2 \langle E_3 E_4 \cos \delta' \rangle \\ U &= 2 \langle E_3 E_4 \sin \delta' \rangle \\ V &= \langle E_3^2 \rangle - \langle E_4^2 \rangle \quad , \end{aligned} \qquad 1.3.1$$

where $E_3(t)$ and $E_4(t)$ are component amplitudes in the left and right-handed antennas and δ' is the phase difference between them. Clearly, the RF hardware shown in figure 1-2 can again be used to derive the parameters. For brevity, discussions and diagrams in this thesis which relate to correlation polarimeters assume the use of antennas sensitive to linear polarization.

In general, to obtain the four Stokes parameters four independent quantities need to be measured. A correlation polarimeter measures the intensities of opposite field components and (effectively) the amplitude and phase of the zero-lag cross-correlation function. If suitable correlation hardware is not available, the Stokes parameters can be formed in an adding polarimeter using intensity signals derived from four antennas. For example, Kraus (1966) shows that using the antennas in figure 1-3, the Stokes parameters can be expressed as:

$$\begin{aligned} I &= I_l + I_r \\ V &= I_l - I_r \\ Q &= 2I_x - (I_l + I_r) \\ U &= 2I_t - (I_l + I_r). \end{aligned} \qquad 1.3.2$$

In reality, it is not necessary to have four separate antennas because the required signals can be synthesized from two antennas as shown in figure 1-4. The phasing network can be placed at the antennas (in which case four receivers are necessary) or at the IF

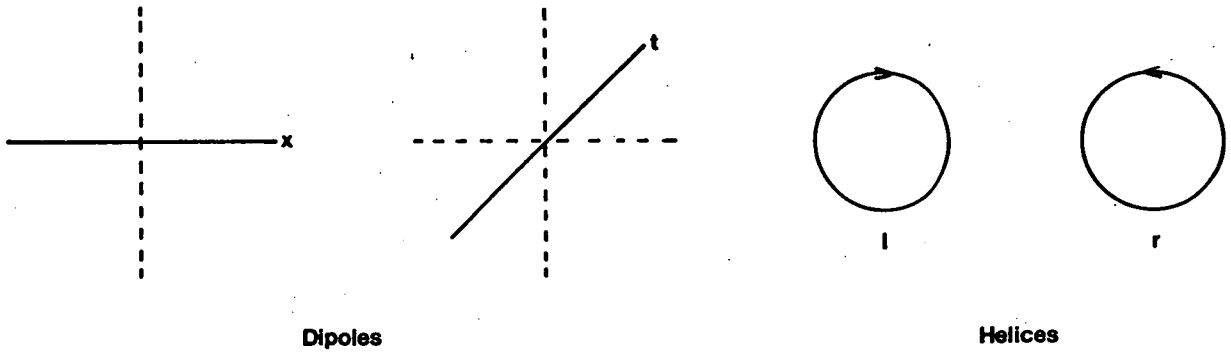


Fig.1-3: Antenna arrangement for deriving Stokes parameters using only intensity measurements.

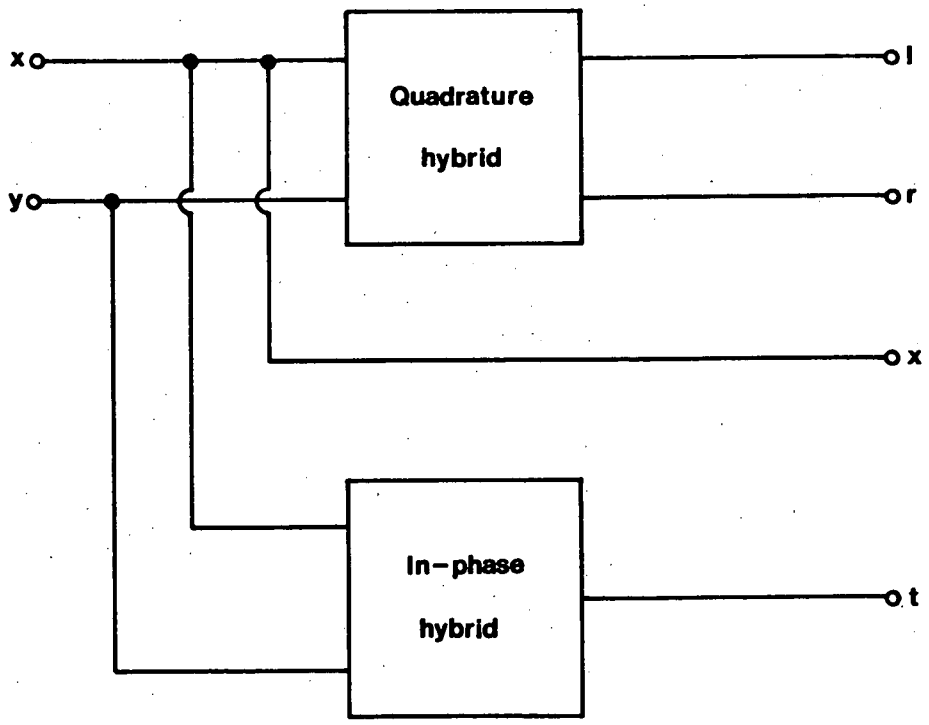


Fig.1-4: Phasing system to simulate the antennas shown in figure 1-3 using signals derived from perpendicular dipole antennas.

outputs if phase coherence is preserved in a two-channel receiving system. Practical instruments are almost always two-channel types.

The disadvantage of an adding polarimeter is that receiver gain errors translate very directly into errors in the output quantities. Three Stokes parameters are formed by differencing signals which typically have a very low signal-to-noise ratio (SNR). Careful calibration and the use of a stable receiving system minimizes errors but the differencing process is inherently unreliable. A correlation polarimeter is preferable because only one parameter is formed by taking differences. A variation in receiver gain still produces a change in the scale of the products but the scale change is more readily removed by calibration procedures than are the multiple changes in an adding polarimeter.

After alignment of a polarimeter, residual amplitude and phase errors can be quantified by applying signals having known amplitude and phase characteristics to the instrument. The response of the polarimeter to these signals can be used as calibration information, allowing corrected Stokes parameters to be derived during data processing. Thiel (1976) describes a mathematical approach to calibration and, after a formal analysis of polarimeter systematic errors as well as the uncertainty in determining the errors, concludes that "polarimeter performance is not arbitrary but should be as good as possible in order not to decrease the accuracy". Stinebring et al. (1984) have demonstrated an accurate, practical polarimeter calibration procedure based on extended observations of pulsar signals.

As well as misalignment, coupling between channels is a common cause of polarimeter error. Cordes and Hankins (1977) have examined the effect of cross-channel coupling and other defects in practical instruments. The major source of coupling is usually imperfect isolation of the two nominally independent antennas. An imperfectly isolated system conserves both total measured power and total polarized power. However, linearly polarized power is seen as circularly polarized power and vice-versa. One means of specifying the interaction between circular and linear polarization channels is by quoting either $L_{spurious}/|V|$ or $|V_{spurious}|/L$. These ratios have limiting values given by $2\sqrt{\eta}$ where η is the magnitude of the cross-polarized power. For coupling levels of -20 dB and -30 dB the limiting values are 0.2 and 0.06 respectively.

1.4 EFFECTS OF PROPAGATION MEDIA ON COSMIC RADIATION

When a polarized electromagnetic wave enters a magnetoionic medium it is propagated within the region in two independent modes which have different phase velocities. If the wave frequency is much greater than the plasma frequency and the electron gyro frequency, and if the magnetic field is not almost exactly transverse to the propagation direction, propagation in the medium is quasi-longitudinal (Ratcliffe, 1962). The two propagation modes are then circularly polarized in opposite senses. If radiation entering the medium is linearly polarized, the resultant of the two natural modes at any point in the medium is also linearly polarized. However, because the modes have different propagation velocities the polarization plane rotates as the wave passes through the medium; this is known as Faraday rotation. For a full summary of Faraday rotation effects see Cohen (1958), the reference on which the initial discussion in this section is based.

When the incident wave is elliptically polarized, the polarization ellipse axial ratio and sense of field vector rotation do not change as the wave propagates but its position angle (ψ in figure 1-1) does. The Faraday rotation a wave experiences in propagating through a magnetoionic median is

$$\theta = (2.36 \times 10^{-3}) f^{-2} \int_0^z n_e B_{\parallel} \, dl \quad \text{radian,} \qquad 1.4.1$$

where f is the wave frequency (expressed in megahertz in this and all subsequent equations in this chapter), n_e is the electron density (cm^{-3}), $B_{\parallel}$ is the longitudinal magnetic field (Gauss) and z is the path length (km).

The angle θ is frequency dependent, so a broadband elliptically polarized signal is dispersed into a continuum of ellipses having a spread in orientation. With a non-zero observing bandwidth dispersed ellipses are added together, resulting in a de-polarization effect. The magnitude of the angular dispersion across the observing band can be deduced from equation 1.4.1 as

$$\gamma \sim 2\left(\frac{\Delta f}{f}\right)\theta \quad \text{radian},$$

if $\Delta f/f << 1$.

From equation 1.4.1, the rotation angle can be expressed as

$$\theta = (RM)\lambda^2 \quad \text{radian}, \qquad 1.4.2$$

where λ is the wavelength (m). The constant (RM) is the rotation measure of the medium (in radian m^{-2}) and is taken as positive when the magnetic field is directed towards the observer. The angular dispersion over a narrow band can be re-expressed as

$$\gamma \sim 2\ RM\left(\frac{300}{f}\right)^2 \frac{\Delta f}{f} \text{ radian}. \qquad 1.4.3$$

Angular dispersion is clearly most severe at low frequencies and is particularly important when considering the effect of Earth's ionosphere on low frequency cosmic radiation. Cohen (1958) cites studies by Hatanaka showing that de-polarization due to angular dispersion becomes significant when $\gamma \sim 1$ radian. Hatanaka's analysis was for a rectangular passband but the general case has been examined by Akabane and Cohen (1961). They show that the correlation factor (equation 1.2.3) measured using a non-zero bandwidth is the product of the true factor (measured using an infinitesimal bandwidth) and the Fourier transform of the polarimeter normalized frequency characteristic.

Hamilton et al. (1977) approximate the ionosphere rotation measure as:

$$RM_i = 0.0814\ (f_0F_2)^2\ B\frac{\cos\delta_B\ \cos(h+h_B) + \sin\delta_B\ \sin\delta}{\cos L\ \cos\delta\ \cosh + \sin L\ \sin\delta} \qquad 1.4.4$$

radian m^{-2}. In this expression f_0F_2 is the ionosphere critical penetration frequency at vertical incidence, h_B and δ_B are the hour angle and declination of Earth's magnetic field direction at the observatory, h and δ are the source hour angle and declination, L is the observatory latitude and B is the field density (Gauss). The RM is negative in the southern hemisphere and the approximation is most accurate for small zenith angles. Even though ionospheric angular

dispersion is less significant at higher observing frequencies, the rotation of absolute position angles is usually of significance below 1 GHz.

Zheleznyakov (1970, ch.2) has summarized the effects of selective absorption of components propagated in different modes in the ionosphere and concludes that non-polarized emission passing through the ionosphere acquires circular polarization. He estimates that the degree of acquired polarization is about 1% at 30 MHz, rising to 10% at 15 MHz and increasing rapidly as the frequency decreases. The effect is negligible above 30 MHz except when accurate metre wavelength polarimetry of weakly polarized sources is undertaken.

Radiation from radio sources outside the solar system is modified by propagation effects in the interstellar medium. The medium is a tenuous plasma ($n_e \sim 0.03$ cm^{-3}) permeated by a weak magnetic field ($B \sim 1$ μG). These conditions result in quasi-longitudinal radio wave propagation within the plasma. Cohen (1960) has estimated selective mode absorption effects in interstellar space to be negligible but dispersion effects are often significant.

Interstellar dispersion is manifested in two forms. Radiation is subjected to Faraday rotation and angular dispersion as already described. In addition, frequency components of impulsive radiation (such as that from pulsars) are delayed differentially in reaching the observer, with high frequencies arriving before lower frequencies. In a finite bandwidth receiver, frequency components at the passband edges are delayed with respect to each other. Thus, the output is not simply the impulse response of the frequency shaping system as it would be for non-dispersed impulsive signals, but rather a response lengthened in time by an amount dependent on the differential delay.

The arrival time for dispersed impulses can be written as

$$t = t_0 + \frac{DM}{2.410 \times 10^{-4}} \frac{1}{f^2} \quad \text{s},$$

where t_0 is the free-space (or light) arrival time and the dispersion measure (DM) is defined by

$$DM = \int_0^z n_e \, dl \ .$$

With the electron density (n_e) in cm^{-3} and the path length (z) in parsecs, DM is expressed in the hybrid units of pc cm^{-3}.

The differential delay between two frequencies is

$$\Delta t_d = \frac{DM}{2.410 \times 10^{-4}} \left[\frac{1}{f_1^{\,2}} - \frac{1}{f_2^{\,2}} \right] \quad s, \qquad 1.4.5$$

with the magnitude of the instantaneous sweep rate being

$$\frac{df}{dt} = \frac{f^3}{8.3 \times 10^3 DM} \qquad MHz\ s^{-1}. \qquad 1.4.6$$

The effect of the very weak interstellar magnetic field is neglected in these expressions, equation 1.4.5 being accurate to about 1 part in 10^4 (Hankins and Rickett, 1975). For pulsar signals a dispersion smear (DS) can be defined such that

$$DS = 100 \frac{\Delta t_d}{P} \quad \%, \qquad 1.4.7$$

where P is the pulsar period in seconds and the passband edge frequencies are used in equation 1.4.5.

Both the time and polarization distorting properties of plasma media are minimised by choosing narrow observing bandwidths. However, radiometer sensitivity scales as $\sqrt{\Delta f \ \Delta t}$ where Δt is the post-detection integration time. Choosing Δf small enough to minimize distortion may reduce the SNR significantly since Δt cannot be increased indefinitely, especially in applications where the fine temporal structure of the signal is important.

In decametric wavelength polarimetry of solar or Jovian bursts for example, Δf is usually chosen to approximate the instantaneous burst bandwidth, with the upper limit on the bandwidth being determined by ionospheric angular dispersion. Bandwidths of 10 or 20 kHz are common. Bandwidths in higher frequency observations are normally determined solely by the radiation characteristics.

For pulsar signals the time resolution (assuming no post-detection time constant) is governed either by the receiver risetime ($\sim 1/\Delta f$) or the dispersion smear time Δt_d. Optimum resolution consistent with best SNR (widest permissible bandwidth) occurs when $1/\Delta f \sim \Delta t_d$ or

$$\Delta f \sim \frac{1}{\Delta t_d} \quad . \qquad 1.4.8$$

For most pulsars this condition gives unrealistically narrow bandwidths from the SNR viewpoint.

If Δf is increased to obtain a higher SNR, the dispersion smear can be calculated from equations 1.4.5 and 1.4.7. Figure 1-5 is a plot of 1% smearing bandwidth as a function of observing frequency and pulsar parameter DM/P. Most pulsars have DM/P < 1000 although DM/P for one recently discovered milli-second pulsar is about 45 000.

If the optimum resolution condition of equation 1.4.8 obtains, the absolute value of the angular dispersion can be expressed as

$$\gamma = 1.976\ f^{-3/2} \left(\frac{RM}{DM}\right) \sqrt{DM} \ \text{radian.} \qquad 1.4.9$$

The ratio RM/DM ~ 1 (μG) and is the approximate mean component of the interstellar magnetic field along the line of sight (Manchester and Taylor, 1977, ch.7). For all pulsars $\gamma << 1$ at observing frequencies above 100 MHz, even with $\Delta f >> 1/\Delta t_d$. Bandwidths can therefore be chosen on the basis of frequency dispersion effects alone.

1.5 SIGNAL REPRESENTATION

Cosmic radio signals may exhibit temporal variations in both intensity (or amplitude) and frequency. As well, the intensity of all sources is frequency dependent. Pictorially, the representation in Figure 1-6 is useful in visualizing the space in which signals are defined. In addition to the three dimensions shown, at least one more dimension is added if even a single wave polarization parameter is represented. Depending on the type of emission, the signal behaviour in any plane (frequency-time, amplitude-frequency etc.) may provide information about the source.

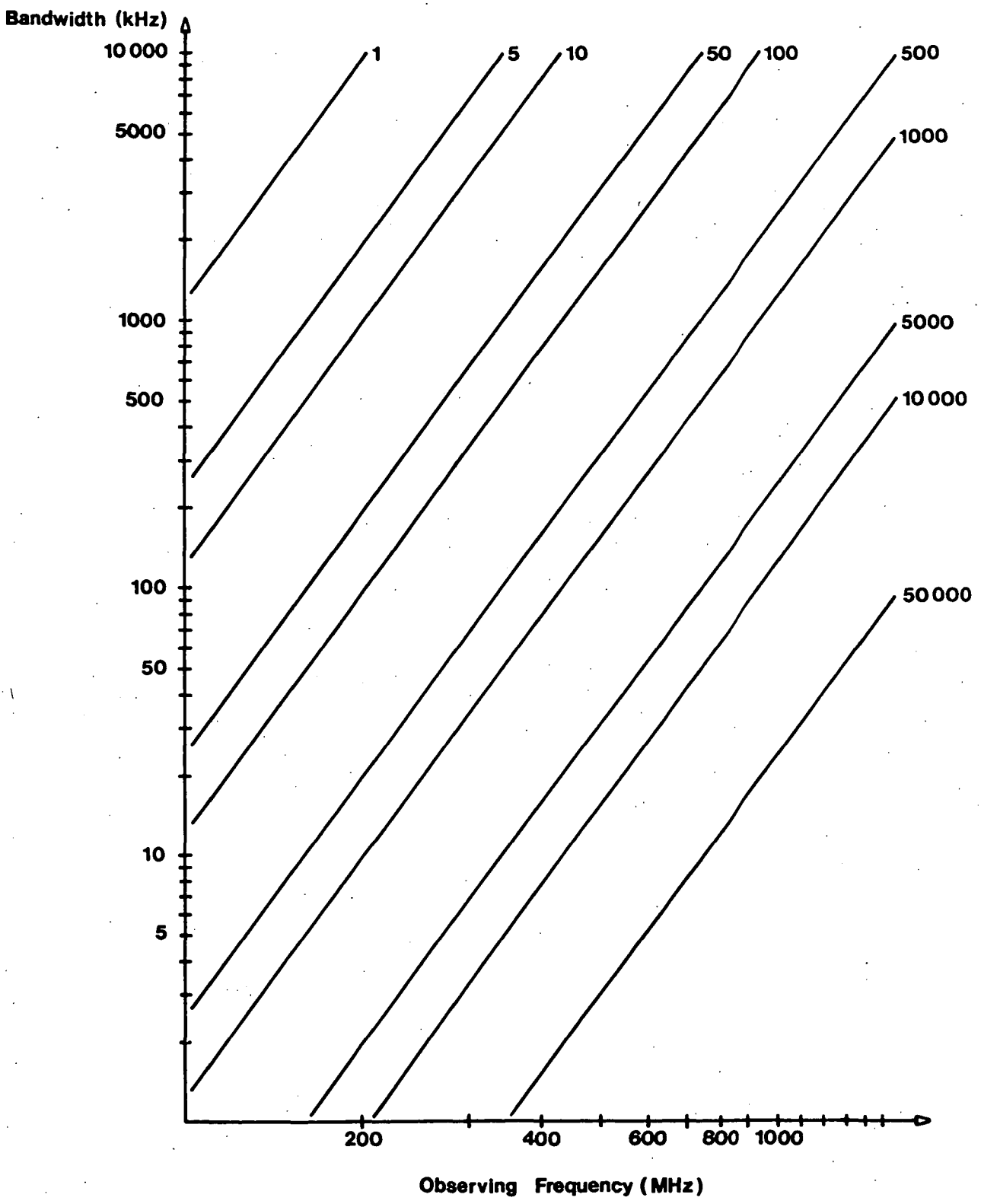


Fig.1-5: Bandwidth for 1% smearing as a function of observing frequency for various values of the pulsar parameter DM/P.

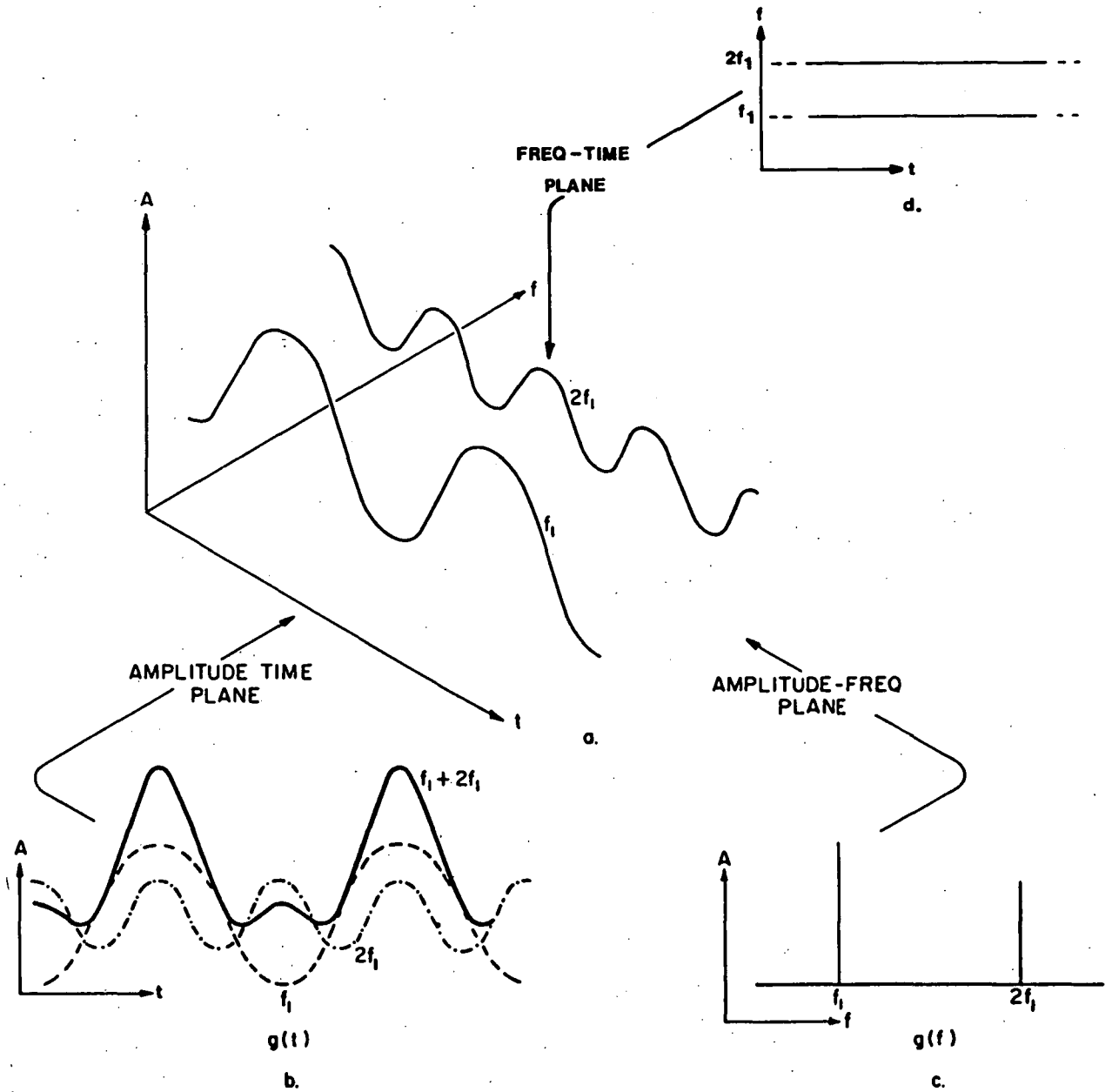


Fig.1-6: Three-dimensional signal representation. The amplitude and frequency properties of actual cosmic radio signals are defined in statistical terms; for clarity deterministic signals are illustrated. Two sinusoids are shown in amplitude-frequency-time space in (a). The amplitude-time plane is shown in (b) and corresponds to the view obtained using an oscilloscope. In (c) the amplitude-frequency plane is shown. This view is obtained using a spectrum analyser. The dynamic spectrum (spectrogram) is shown in (d) and is the signal projection in the frequency-time plane. (Adapted from Hewlett-Packard Application Note No.63).

Ideally, observations would allow reconstruction of the entire data field, subject to the uncertainty relationship between time and frequency resolutions. In practice, a multi-dimensional representation is rarely reconstructed. In some cases signal characteristics in a particular plane yield little information. Even when full reconstruction is desirable it is rarely attempted because SNR constraints usually make it necessary to average over one or more signal domains. Before this sensitivity limit is reached however, more mundane restrictions such as technological and economic limitations on observing hardware are often significant.

Full representation of signals clearly involves spectral information. When conventional spectrometers are used it is difficult to obtain the high time resolution coherent spectra needed for polarization studies of non-stationary signals (chapter two). The SAW polarimeter spectrometer is the result of a project aimed at providing astronomers with an instrument able to reconstruct as fully as possible the multi-dimensional signal representation. In particular applications the observer is free to average or discard information from one or more dimensions to increase the SNR or to minimise data processing requirements.

In the following sections a few topical astrophysical applications of a polarimeter spectrometer are mentioned. Dynamic emission from Jupiter, the sun and pulsars is considered in some detail since these signals are particularly likely to yield new information if full characterization is possible. However, even continuum source observations are enhanced by the use of a polarimeter spectrometer as the following example illustrates.

Goldstein and Reed (1984) used a multi-channel receiving system in conjunction with a rotatable feed to study the radio source 3C27, assumed to be a radio galaxy. This source has major components separated by 45" and it is shown that an anomalous position angle versus frequency curve is well-explained if the components have separate rotation measures of 165 and -104. The sign change indicates that the net magnetic field reverses between the components, perhaps because a large field in the central part of the source covers one region and not the other.

A more intriguing suggestion is that our own galaxy contains a dipole field with a scale of about 1 pc. One component of the source is seen inside the current loop associated with the field; the other is outside the loop. Goldstein and Reed suggest the likelihood of more loops in the Milky Way and speculate on possible locations. A polarimeter spectrometer would be of great value in the search for these galactic current loops since position angle data are available directly, greatly simplifying data acquisition and reduction techniques.

1.6 DECAMETRIC JOVIAN RADIATION

The observed characteristics of various burst emissions from Jupiter have been described in detail by Ellis (1974). Current theories of generation and propagation are outlined by several authors in the work edited by Dessler (1983). The burst emission is basically electron cyclotron radiation and is generated in the range 10 kHz - 40 MHz. The lower limit has been observed by interplanetary probes and corresponds to the maximum plasma frequency between the source and the spacecraft. The upper limit is the highest cyclotron frequency in the Jovian magnetosphere. The intensity of the decametric radiation is correlated strongly with the position of Io (the inner Galilean satellite) and with the orientation of Jupiter with respect to the observer.

Bursts are characterized by their duration, frequency span, slope, and instantaneous bandwidth in the frequency-time plane. High time resolution studies have yielded large amounts of information about several source regions in the magnetosphere. Nevertheless, Carr et al. (1983) point out that complex phenomenology, difficulties in separating first and second order effects, and the presence of propagation effects and observational artifacts have all been factors preventing location of the actual radio sources.

The decametric radiation from Jupiter has clear circular polarization properties which have been related to cyclotron emitters in the Jovian tilted-dipole magnetic field. An alternative view is that the polarization properties arise because of propagation effects in plasma media near the source. Carr and Desch (1976) conclude that spectra having high time and frequency resolutions provide a wealth of

detail related to emission and propagation processes and that persistent study is warranted.

Recent observations of the Jovian magnetosphere by interplanetary probes have also led to the conclusion that further ground-based studies are necessary. For example, Alexander and Desch (1984) have examined Voyager 1 and 2 observations of the milli-second "S" bursts. They conclude that although the bursts exhibit significant circular polarization, the polarization sense is not well explained by existing models of the radiation and propagation processes. They stress that more ground-based polarization data are essential.

Ground-based decametric wavelength polarimetry is difficult because of ionospheric rotation, angular dispersion and scintillation. Polarization investigations to date have often been ambiguous in outcome but have demonstrated that all four Stokes parameters must be recorded if the study is to be effective. Two high gain, oppositely polarized antennas are therefore required. The gain requirement arises because burst data are analysed at high frequency and time resolutions, so high SNR is needed for clear results.

In fact, 1985 is an optimum time to conduct a ground-based Jovian polarimetry program in Tasmania. Jupiter has reached its most southern declination, so the zenith angle is about 23^{o} and the declination at transit changes quite slowly over a period of a year or so. A sensitive, fixed antenna array would therefore remain oriented approximately correctly for some time. Since the zenith angle is fairly small, it may be possible to find a natural slope on which to construct the antenna. Ionosphere critical frequencies are often low in Tasmania because the geomagnetic latitude is high and this, together with the approaching solar minimum, should lead to minimal ionospheric influence. It is reasonably certain that a wideband, high time resolution polarimeter spectrometer ($\lesssim$ 30 kHz resolution bandwidth) operated under these conditions would produce valuable data, the broadband nature of the instrument helping to identify propagation artifacts. The dual-channel SAW spectrum analysis system described in appendix five could be used as a polarimeter spectrometer with parameters particularly suitable for observations of Jovian burst emission.

1.7 SOLAR RADIO BURSTS

A rich variety of non-stationary solar radiation has been observed over more than three decades. Characteristics of the radiation have been summarized by Zheleznyakov (1970, ch.4) and, more recently, in monographs by Elgaroy (1977) and Kruger (1979).

Burst signals can occur over wide regions of the radio spectrum and last from milli-seconds to minutes. Burst occurrence is linked statistically with visible solar flares although the connection is not firm for individual events. The transient emission is divided into two broad categories: microwave bursts and decametric bursts. As with Jovian bursts, individual events in each category are classified by their appearance in the frequency-time plane.

Kruger (1979, ch.3) has characterized microwave bursts ($f \gtrsim 1$ GHz) as "short duration enhancements of the flux level when the whole-sun flux density is monitored by single-frequency radiometer observations". Both the spectral and temporal behaviour of microwave bursts is much smoother than that of the decametric types. Microwave bursts are, in general, partially circularly polarized with the degree of polarization tending to increase at higher frequencies. The sense of polarization is usually explained by considering the orientation of escaping rays relative to solar magnetic field lines. Theories of microwave burst generation involve both bremstrahlung and gyro-synchrotron emission processes and are discussed in great detail by Zheleznyakov (1970, ch.9).

Decametric bursts often exhibit very complex structure in both the time and frequency domains. Bursts are loosely classed as fast or slow-drift types, depending upon their slope in the frequency-time plane. Elgaroy (1977, ch.8 et. seq.) lists the characteristics of many burst types and discusses possible generation mechanisms. Decametric bursts usually exhibit significant circular polarization and there are several reports of a linearly polarized component in some bursts at certain times (Komesaroff, 1958; Akabane and Cohen, 1961; Bhonsle and McNarry, 1964; Harvey and McNarry, 1970; Chin et al., 1971; Dodge, 1972). However, studies by Grognard and McLean (1973) and Boischot and Lecacheux (1975) found no evidence to confirm claims of linear polarization in fast-drift type III bursts.

Present emission models do not allow the existence of linear polarization. Elgaroy (1977, ch.3) points out that, for all likely source dimensions, linear polarization should be destroyed by angular dispersion in the source region for $\Delta f \gtrsim 100$ Hz. McConnell (1981) notes that the presence of linear polarization could be explained if the polarizer were higher in the corona than the source. This is not implausible as both Cohen (1960) and Melrose (1975) have suggested theoretical coronal mechanisms for inducing linear polarization. An indisputable indication of linear polarization in solar bursts would be of great importance since it would give essential information about emission and propagation mechanisms and about the coronal magnetic field.

At present, most solar astronomers reject reports of linear polarization and cite defects (such as ground reflections) in polarimetry systems as the likely origin of the observed linearly polarized component. Nevertheless, the wide implications of any reliable discovery of a linearly polarized fraction make the question worthy of further examination. In particular, the usual practice of recording only two intensity signals derived from opposite-hand circular antennas is unsound since the chance of identifying sporadic linear polarization is zero.

The dual-channel SAW spectrum analyser described in appendix five has higher frequency resolution than receivers used by many investigators reporting a linearly polarized component in decametric bursts. When used as a polarimeter spectrometer, this instrument has the potential to yield results which would be a valuable contribution to the debate. As with Jovian observations, the wideband nature of the instrument should assist in identifying spurious propagation effects.

1.8 PULSAR RADIATION

Since the announcement by Hewish et al. (1968) of the detection of a pulsating radio source, over 330 such objects have been found. The pulsation periods range from milli-seconds to seconds and all but one of the known pulsars are in our own galaxy. The characteristics of pulsars, together with several current emission models, are summarized in monographs by Smith (1977) and Manchester and Taylor (1977).

The most significant radiation features are the extreme pulse regularity, the intense broadband nature of the low duty cycle pulsed signal, and the similar polarization properties of many objects.

Pulsars are believed to be rotating, magnetized neutron stars. The radio emission is supposed to originate from a directional source located in a co-rotating magnetosphere. Rotation of the star may cause the radiation beam to intersect the observer's line of sight, resulting in the reception of a pulse each star revolution. The time-scale of pulsar signals is often expressed in rotation longitude with one pulse period corresponding to 360^{o}. The great rotational inertia of the neutron star provides a natural explanation for the pulse regularity; the emission intensity (brightness temperature $\sim 10^{30}$ K) is explained by invoking coherent source models or maser amplification mechanisms external to the source. Spectral and polarization properties are accounted for in different ways depending upon the emission model used.

Various models place the emission region outside the light cylinder (the imaginary surface in space at which a particle co-rotating with the star would attain the velocity of light), in the region of closed magnetic field lines near the light cylinder, or along open magnetic field lines near the star's magnetic poles. At present, the last (polar cap) models are the most developed.

Pulsar signals seem to reveal progressively more information as the completeness of the multi-dimensional signal representation (section 1.5) is increased. In general, reconstruction is hampered by the low SNR obtainable when observing the objects. Also, combined spectral and polarization measurements are rarely made because of the hardware complexity when conventional spectrometers are used.

For many pulsars an integrated profile (the synchronous average of many pulses) is the only characterization available, individual pulses being either too weak to detect or observable only with low SNR. The integrated intensity profile of a given pulsar is usually fairly stable and integrated polarization measurements show similar stability. Most pulsars exhibit significant linear polarization, the position angle of which changes across the pulse. Circular polarization of

either sense may also be present but is usually weak compared with the linear component.

The gross properties of the position angle change are often accounted for using a single vector model first proposed by Radhakrishnan and Cook (1969). Depending on the alignment between the line of sight and the pulsar's magnetic dipole axis, smooth, monotonic position angle swings of up to 180^{o} are predicted. Many pulsars show behaviour consistent with the model. However, in some pulsars the position angle is subject to one or more orthogonal transitions (flips) in which the angle changes by up to 90^{o} over a very small longitude interval. The phenomonen is described by Backer et al. (1976) who propose a geometrical explanation of the behaviour in one pulsar.

The presence of circular polarization in pulsar radiation is not easily accounted for. In polar cap emission models, radiating particles accelerated along polar magnetic field lines are probably highly relativistic because of magnetospheric electric field conditions. Their radiation is therefore assumed to be from curvature or synchrotron mechanisms. Cordes and Hankins (1977) note that such radiation from an ensemble of particles is unlikely to possess significant circular polarization because of cancellation effects in the emission region.

Cocke and Pacholczyk (1976) and Melrose (1979) have proposed mechanisms by which linear polarization could be partly converted into elliptical polarization by propagation phenomena. These models also predict position angle flips under certain propagation and magnetospheric plasma conditions. However, attempts to deduce realistic conversion parameters using actual pulsar data have not been successful (Cocke and Pacholczyk, 1980). Melrose (1979) suggests that studies of the spectrum of polarization flips be undertaken in an attempt to verify the prediction of a fairly strong frequency dependence in polarization conversion processes.

Studies of individual pulses from strong pulsars show that a great variety of short duration polarization effects occur. Individual pulses or sub-pulses (narrow intensity peaks within the pulse) are often very nearly completely polarized and orthogonal flips may occur within a sub-pulse. The sense of circular polarization

sometimes reverses at the transition longitude but in at least one case the sign of circular polarization changes within a sub-pulse with no apparent discontinuity in the position angle of the linear component (Cheng and Ruderman, 1979).

Single pulse polarization data are available for fewer than 40 pulsars, the great majority of which are northern-sky objects. Two of the larger surveys are those of Manchester et al. (1975) and Stinebring et al. (1984). The most important point to emerge from these studies is that synchronous averaging destroys a large amount of information about the emission process, often washing out important pulse-to-pulse polarization details. Even single pulse observations with moderate to high time resolution may not provide full information. Cordes and Hankins (1977) studied the 430 MHz polarization micro-structure of four strong pulsars and showed that micro-pulses (features with a time-scale of micro-seconds) are even more highly polarized than sub-pulses, with orthogonal flips often occurring at micro-pulse edges.

Bjornsson (1984) postulates that pulsar polarization is a direct consequence of the emission process. His model predicts position angle flips at the sub-pulse and micro-pulse level under certain conditions but unless the notion of separate sub-pulse and micro-pulse origins is accepted, it is still necessary to suppose that propagation effects cause circular polarization. Bjornsson cites studies suggesting that linearly polarized features are more highly correlated with respect to frequency than are circularly polarized features. This, together with the rather weak correlation between circularly and linearly polarized transitions at a given frequency, is evidence in favour of propagation induced circular polarization.

Systematic spectral and polarization studies would be of great value in examining further the frequency correlation of short time-scale polarization structure. Stinebring et al. (1984) used two narrowband polarimeters separated in frequency to observe 11 strong, low dispersion pulsars. They conclude that the frequency dependence of polarization phenomena is, in general, weak. However, the small sample of pulsars in their investigation makes further study necessary.

The evidence that more complete reconstruction of a multi-dimensional representation of pulsar signals leads to identification of new phenomena was the primary motivation for the development of the high time resolution polarimeter spectrometer described in this thesis. It is envisaged that the new instrument and its associated equipment will have four main observational modes. Firstly, for strong pulsars single pulses are visible in individual frequency resolution cells (~650 kHz), so spectral and polarization characteristics of single pulses are obtainable over a wide bandwidth (~30 MHz). With common data acquisition computers a time resolution of about 150 μs is feasible even if full Stokes polarimetry is undertaken. With dedicated back-end processors resolutions down to the 2 μs SAW spectrometer limit are possible, although dispersion smearing across a frequency cell means that this order of resolution would be useful only at the highest observing frequencies. For example, at 650 MHz and 2 GHz the dispersion smear times are:

$$\Delta t_d\ (650) = 19.6\ DM \quad \mu s$$

and 1.8.1

$$\Delta t_d\ (2000) = 0.67\ DM \quad \mu s.$$

Secondly, the instrument will provide individual pulse data on many pulsars which are too weak to be observed using narrowband polarimeters. In this mode, pulses from different frequency cells are averaged in the frequency domain to yield one pulse with a high SNR. Pulses must of course be combined such that the effects of frequency dispersion and Faraday rotation are accounted for. The correction processes for these effects are known as de-dispersion and de-rotation.

In the third mode, pulses are integrated synchronously in each frequency cell to yield integrated polarization profiles as a function of frequency. Finally, for the weakest pulsars, integrated profiles from each cell are combined to give a single profile. Thus, the instrument permits very full reconstruction of the pulsar signal, while giving the observer the option of any degree of averaging in the time or frequency domains should this be necessary to maintain

an acceptable SNR. The operational modes outlined are basic ones and many other methods of signal processing can be envisaged for particular experiments.

In the next chapter several spectrometer types are examined and the suitability of each as the basis of a high time resolution coherent processor or polarimeter is assessed. Surface acoustic wave spectrometers turn out to be well-suited to this role and the principles of SAW spectrum analysis are discussed prior to a more detailed description in later chapters.

CHAPTER 1 - REFERENCES

Akabane, K. and Cohen, M.H.: Polarization measurements of type III bursts and Faraday rotation in the corona. Astrophys. J., Vol.133, 1961, pp.258-268.

Alexander, J.K. and Desch, M.D.: Voyager observations of Jovian millisecond radio bursts. J. Geophys. Res., Vol.89, No.A5, 1984, pp.2689-2697.

Backer, D.C., Rankin, J.M. and Campbell, D.B.: Orthogonal mode emission in geometric models of pulsar polarization. Nature, Vol.263, 1976, pp.202-207.

Bhonsle, R.V. and McNarry, L.R.: Polarization characteristics of type III solar radio bursts at 74 Mc/s. Astrophys. J., Vol.139, 1964, pp.1312-1327.

Bjornsson, C.: Pulsar polarization as a direct consequence of the emission process. MNRAS, Vol.208, 1984, pp.293-307.

Boischot, A. and Lecacheux, A.: Linear polarization in meter and decameter solar radio bursts. Astron. Astrophys., Vol.40, 1975, pp.55-61.

Born, M. and Wolf, E.: "Principles of optics". Ch. 1 and 10. Pergamon, 1970.

Carr, T.D. and Desch, M.D.: Recent decametric and hectometric observations of Jupiter - in "Jupiter" (T.Gehrels ed.). University of Arizona Press, 1976.

Carr, T.D., Desch, M.D. and Alexander, J.K.: Phenomenology of magnetospheric radio emission - in "Physics of the Jovian magnetosphere" (A.J. Dessler ed.). Ch.7, Cambridge University Press, 1983.

Cheng, A.F. and Ruderman, M.A.: A theory of subpulse polarization patterns from radio pulsars. Astrophys. J., Vol.229, 1979, pp.348-360.

Chin, Y.C., Lusignan, B.B. and Fung, P.C.W.: Polarization measurements of solar type III radio bursts at 23.5 MHz. Solar Phys., Vol.16, 1971, pp.135-151.

Cocke, W.J. and Pacholczyk, A.G.: Theory of the polarization of pulsar radiation. Astrophys. J. (Lett.), Vol.204, 1976, pp.13-15.

Cocke, W.J. and Pacholczyk, A.G.: Individual pulse-polarization patterns and the quasi-transverse propagation theory in pulsars. Astrophys. J., Vol.235, 1980, pp.196-198.

Cohen, M.H.: Radio astronomy polarization measurements. Proc. IRE, Vol.46, Jan. 1958, pp.172-183.

Cohen, M.H.: Magnetoionic mode coupling at high frequencies. Astrophys. J., Vol.131, 1960, pp.664-680.

Cordes, J.M. and Hankins, T.H.: Pulsar polarization fluctuations at 430 MHz with microsecond time resolution. Astrophys. J., Vol.218, pp.484-503, 1977.

Dessler, A.J. (ed.): "Physics of the Jovian magnetosphere". Cambridge University Press, 1983.

Dodge, J.C.: Ph.D. thesis. University of Colorado, 1972.

Elgaroy, E.O.: "Solar noise storms". Pergamon, 1977.

Ellis, G.R.A.: The Jupiter radio bursts. Proc. Astron. Soc. Aust., Vol.2, No.5, Sept. 1974, pp.236-243.

Goldstein, S.J. and Reed, J.A.: Double Faraday rotation toward 3C27. Astrophys. J., Vol.283, 1984, pp.540-545.

Grognard, R.J.M. and McLean, D.J.: Non-existence of linear polarization in type III solar bursts at 80 MHz. Solar Phys., Vol.29, 1973, pp.149-161.

Hamilton, P.A., McCulloch, P.M., Ables, J.G. and Komesaroff, M.M.: Polarization characteristics of southern pulsars - I. MNRAS, Vol.180, 1977, pp.1-8.

Hankins, T.H. and Rickett, B.J.: Pulsar signal processing - in "Methods in Computational Physics". Vol.14, Academic Press, 1975.

Harvey, G.A. and McNarry, L.R.: The polarization of solar radio emission at 74 MHz: May 18-26, 1967. Solar Phys., Vol.11, 1970, pp.467-496.

Hewish, A., Bell, S.J., Pilkington, J.D.H., Scott, P.F. and Collins, R.A.: Observation of a rapidly pulsating radio source. Nature, Vol.217, 1968, pp.709-713.

Ko, H.C.: On the reception of quasi-monochromatic partially polarized radio waves. Proc. IRE, Vol.50, Sept. 1962, pp.1950-1957.

Komesaroff, M.: Polarization measurements of three spectral types of solar radio bursts. Aust. J. Phys., Vol.11, 1958, pp.201-214.

Kraus, J.D.: "Radio astronomy". Ch.4, McGraw-Hill, 1966.

Kruger, A.: "Introduction to solar radio astronomy and radio physics". Reidel, 1979.

Longhurst, R.S.: "Geometrical and physical optics". Ch.22, Longman.

Manchester, R.N. and Taylor, J.H.: "Pulsars". Freeman, 1977.

Manchester, R.N., Taylor, J.H. and Huguenin, G.R.: Observations of pulsar radio emission - II. Astrophys. J., Vol.196, 1975, pp.83-102.

McConnell, D.: Ph.D. thesis. University of Tasmania, 1981.

Melrose, D.B.: Small scale inhomgeneities in the solar corona. Solar Phys., Vol.43, 1975, pp.79-86.

Melrose, D.B.: Propagation effects on the polarization of pulsar radio emission. Aust. J. Phys., Vol.32, 1979, pp.61-70.

Radhakrishnan, V. and Cooke, D.J.: Magnetic poles and the polarization structure of pulsar radiation. Astrophys. Lett., Vol.3, 1969, pp.225-229.

Ratcliffe, J.A.: "The magneto-ionic theory and its application to the ionosphere". Ch.8, Cambridge University Press, 1962.

Smith, F.G.: "Pulsars". Cambridge University Press, 1977.

Stinebring, D.R., Cordes, J.M., Weisberg, J.M., Rankin, J.M. and Boriakoff, V.: Pulsar polarization fluctuations - I and II. Astrophys. J. (Supp.), Vol.55, 1984, pp.247-277 and 279-288.

Thiel, M.A.: Error calculation of polarization measurements. J. Opt. Soc. Am., Vol.66, No.1, 1976, pp.65-67.

Wolf, E.: Coherence properties of partially polarized electromagnetic radiation. Nuovo Cimento, Vol.13, Sept.1959, pp.1165-1181.

Zheleznyakov, V.V.: "Radio emission of the sun and planets". Pergamon, 1970.

CHAPTER 2 REAL-TIME SPECTRUM ANALYSIS

2.1 INTRODUCTION

In this chapter several methods of spectrum analysis are reviewed and the suitability of each as the basis of a high time resolution coherent spectrometer is examined. Except for a review of one-bit correlation methods the discussion is confined to real-time, linear processing techniques applicable to the analysis of general signals about which nothing is known or assumed. A real-time analyser is defined as one in which the time required to calculate a spectrum is no greater than that needed to acquire the data to be transformed.

The reciprocal spreading property of Fourier transformation (see, for example, Stuart, 1961) leads to the relationship $\Delta f \Delta t \sim 1$ where Δf and Δt are resolutions in frequency and time. Hence, for a resolution of Δf, a data window having a minimum duration of about $1/\Delta f$ needs to be analysed. The non-linear maximum entropy method of spectrum analysis attempts to circumvent the spreading relationship and has been demonstrated to be effective in at least some astronomical applications. However, the method offers no practical advantages so far as high time resolution analysis is concerned and is not included in this review.

In assessing the relative time resolution performance of each spectrum analyser it is worth noting that in applications where spectral data are transferred to a computer, the time taken to pass data to the computer may exceed any intrinsic analyser speed limitation. For example, a typical general purpose computer requires 1 μs to acquire one data word from a high speed interface. The best time resolution on a 1K-point spectrum is therefore around 1 ms even if no operations other than those involved in the data transfer are undertaken.

If sensitivity is not a prime consideration a fast analyser can burst sample a spectral band. In this mode the spectrometer produces a spectrum having the maximum possible time resolution then ceases operation until data are assimilated by the computer. In most astronomical applications the combined speed limitations of the spectrometer and computer must be considered when estimating time resolution since burst sampling is usually inappropriate, mainly on the grounds of reduced sensitivity. Nevertheless, spectrometer

limitations alone are of interest because some scientific objectives can be met by transferring output from a real-time analyser to a dense, high speed storage medium (such as videotape or moving photographic film) for off-line analysis.

Although full details of pulse compression (SAW) spectrum analysis are given in chapter four, section 2.6 gives an intuitive explanation of the process. To clarify the place of compression analysers in the general area of spectrum analysis figure 2-1 shows the division of the field on the basis of analyser operating principle.

2.2 FILTERBANK SPECTRUM ANALYSERS

The filterbank spectrometer is conceptually the simplest type of spectrum analyser. In the form used in radioastronomy it consists of a receiver with a bank of frequency contiguous IF filter channels (figure 2-2). The IF strips normally include an amplifier to overcome signal splitting and filter transmission losses. Individual IF gain controls allow the overall spectrometer frequency response to be made flat (gain levelled). Each channel output is square-law detected, passed through a post-detection lowpass filter (or integrator) and then made available to signal processing equipment outside the spectrometer. In principle, two filterbanks could be used as the basis of a multi-channel correlation polarimeter (figure 2-3) or other coherent processor.

The frequency resolution of a filterbank analyser is determined by the bandwidth of the individual channels and the uncertainty (standard deviation) of the spectral estimate in a single channel is given by the radiometer equation

$$\sigma = \frac{1}{\sqrt{\Delta f \Delta t}} , \qquad 2.2.1$$

where Δf is the equivalent bandwidth of the IF filter (Tiuri, 1964) and Δt is the post-detection integration time (2RC for an R-C lowpass network). In this and subsequent references to the radiometer equation the mean integrator output is normalized to unity unless otherwise stated.

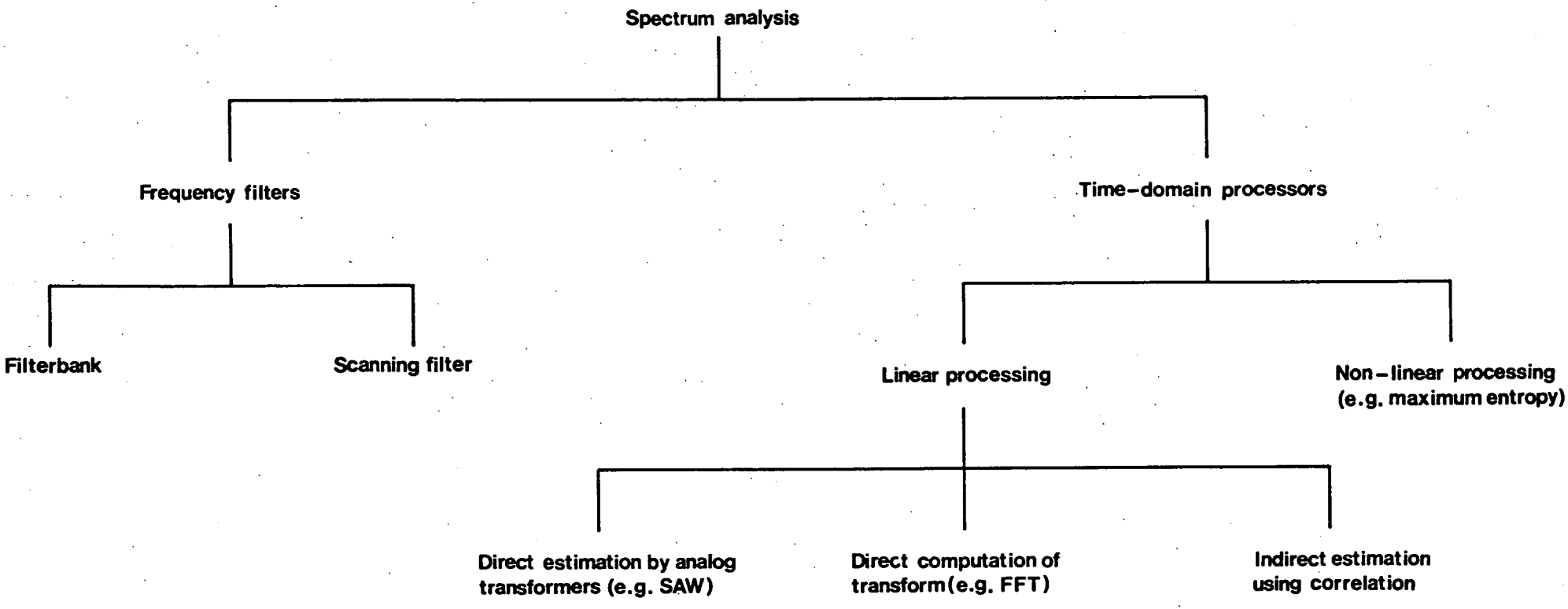


Fig.2-1: Methods of spectrum analysis. (Adapted from Roberts et al., 1980).

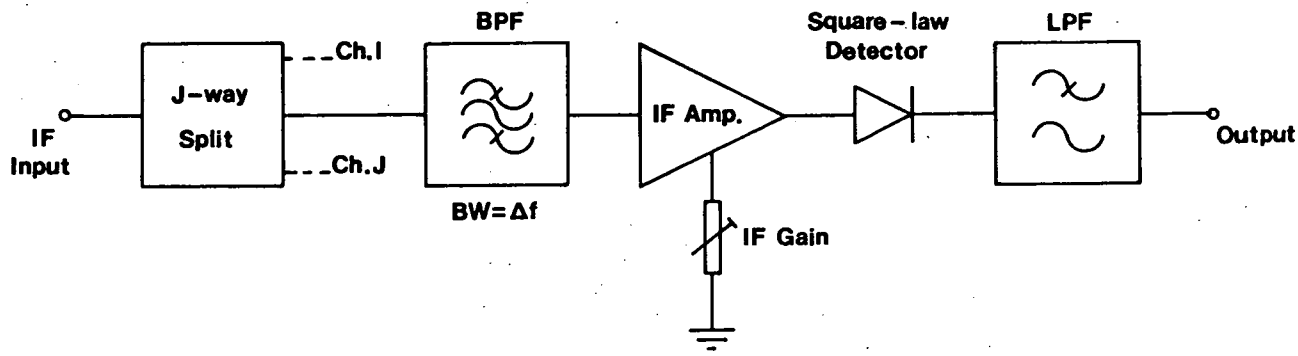


Fig.2-2: Basic form of a J-channel filterbank spectrometer.

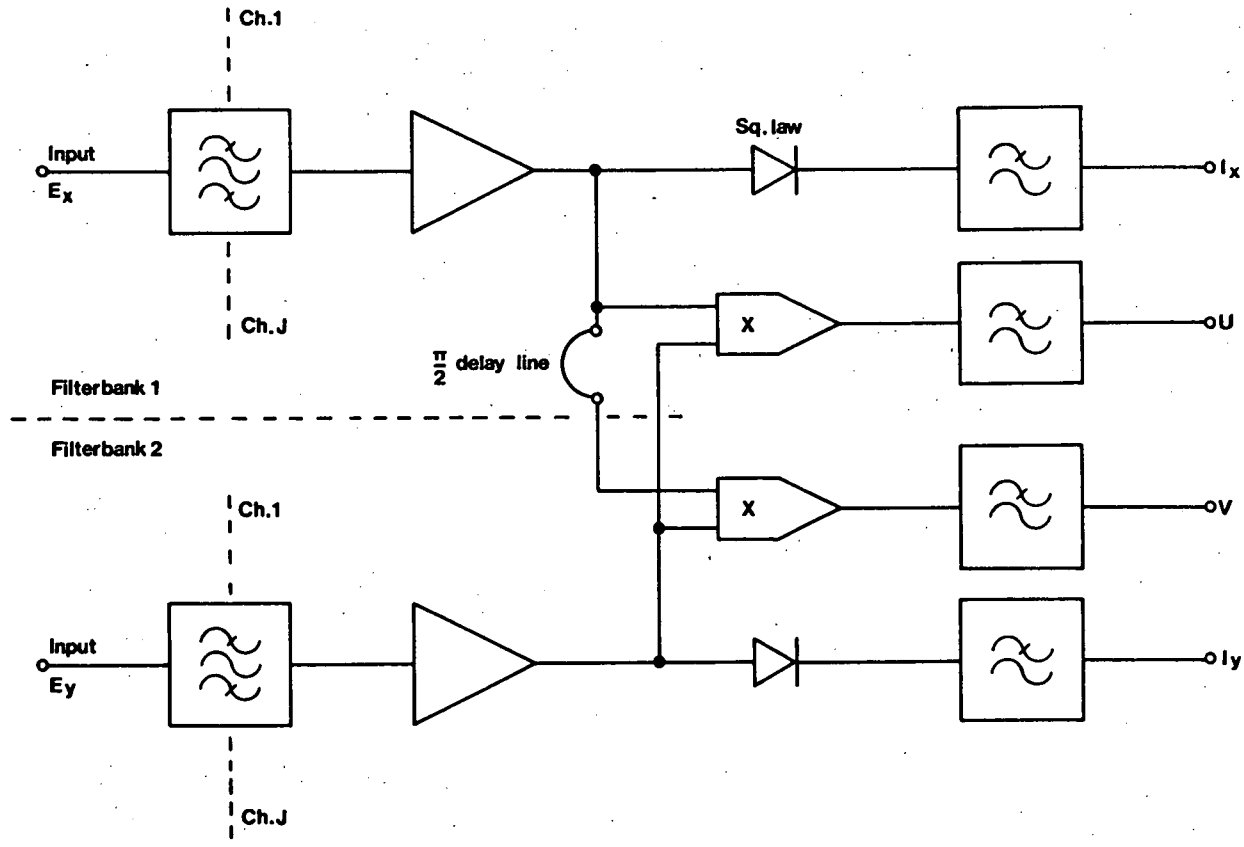


Fig.2-3: Filterbank correlation Stokes polarimeter.

Practical difficulties arise which prevent filterbank analysers being useful in coherent spectrometers when more than a few frequency channels are required. The most obvious objection is that the physical complexity of the spectrometer and output processing stages scales directly as the number of channels. In a J-channel polarimeter for example, 2J IF strips, 2J square-law detectors and 2J high frequency correlators are required (figure 2-3). Each channel must be aligned not only to match the shape factor and gain of its J-1 adjacent-frequency counterparts but also to match the gain, phase and shape characteristics of its same-frequency counterpart in the second filterbank. The whole system must remain gain and phase stable over an extended period of time. Allowing for meticulous design and alignment, the stability specifications are difficult to meet and it is likely that microprocessor controlled gain and phase adjustment of each channel would be required, probably in response to a broadband calibration signal. Such a system has been implemented to gain level a 256-channel filterbank (Ellis, 1984) but the provision of stable, electronically variable phase shift elements in each channel is difficult.

A further problem arises when a large number of channels is needed. Economic considerations usually mean that the IF filters are of an L-C type having at least a two-pole response. It is difficult to build L-C filters with a matching shape factor and roughly the same insertion loss over a very wide bandwidth. For bandwidths exceeding about 20 MHz the total RF bandwidth is normally split into sub-bands. Each sub-band is mixed to the same IF centre frequency and the same basic IF filter set duplicated as many times as is necessary to allow coverage of the required total bandwidth. In a coherent processor, two complete mixing and associated RF and IF filtering schemes are necessary and, as well as complicating construction, it is expected that the proliferation of filters and amplifiers would make amplitude and phase alignment of the instrument difficult.

For a small scale instrument (perhaps 10 channels or less) covering a total bandwidth of less than about 20 MHz, a filterbank coherent processor is practical and given the limitations and complexities of the other approaches to be described, is the best choice.

For incoherent processing (power spectra only) the filterbank approach remains viable with up to several hundred channels. When wideband, high frequency resolution analysis is needed, the cost and physical complexity of filterbanks are major problems and difficulties associated with alignment and alignment stability become acute.

2.3 THE ACOUSTO-OPTICAL SPECTRUM ANALYSER

The acousto-optical spectrometer (AOS) is based on the interaction between RF sound waves propagating in a translucent crystal and laser light passing through the crystal. Figure 2-4 shows the most common arrangement, known as a bulk wave or Bragg cell AOS. RF energy is applied to the Bragg cell transducer and, by virtue of the piezo-electric characteristic of the crystalline solid, produces a bulk acoustic wave within the cell. Viewed from the approach of classical physics, the acoustic wave produces a variation in the refractive index of the propagation medium the form of which is approximately sinusoidal for a CW RF input. The variation in refractive index varies the propagation time of laser light through the crystal, causing the acoustic wave to behave as a moving phase grating which diffracts the light. The diffraction angle is proportional to the grating spacing (determined by the RF signal frequency since the acoustic wave velocity is fixed) and inversely proportional to the laser frequency. At a critical angle of incidence, known as the Bragg angle, the energy in the first order diffraction pattern is a maximum. Fundamentally therefore, the Bragg cell produces an angular variation which is proportional to the RF input frequency.

Light emerging from the Bragg cell passes through a lens which focuses the beam onto a charge-coupled device (CCD) photodetector array. More precisely, the lens produces a Fourier transform of the input angular spectrum, giving a linear displacement in its back focal plane proportional to the object plane angular position. The overall effect is such that varying the RF input frequency produces a change in the linear position of light incident on the photodetector array. Since diffracted light amplitude is proportional to grating (signal) amplitude and position on the detector array is proportional to the signal frequency, the instrument behaves as a spectrum analyser. An output spectrum can be obtained by examining sequentially the charge on each detector cell.

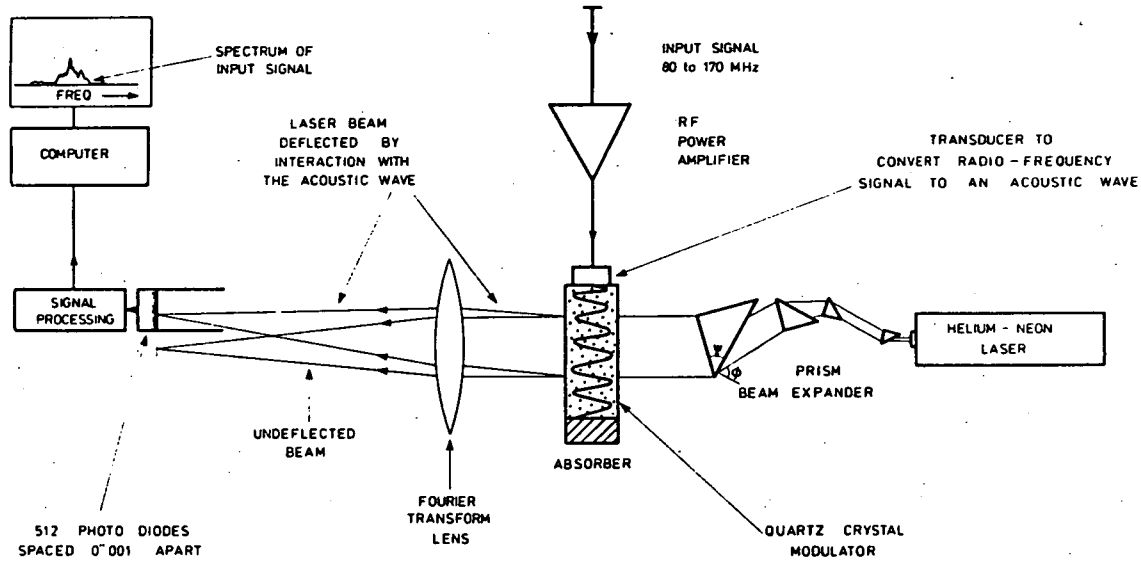


Fig.2-4: CSIRO acousto-optical spectrograph. (After Milne and Cole, 1979).

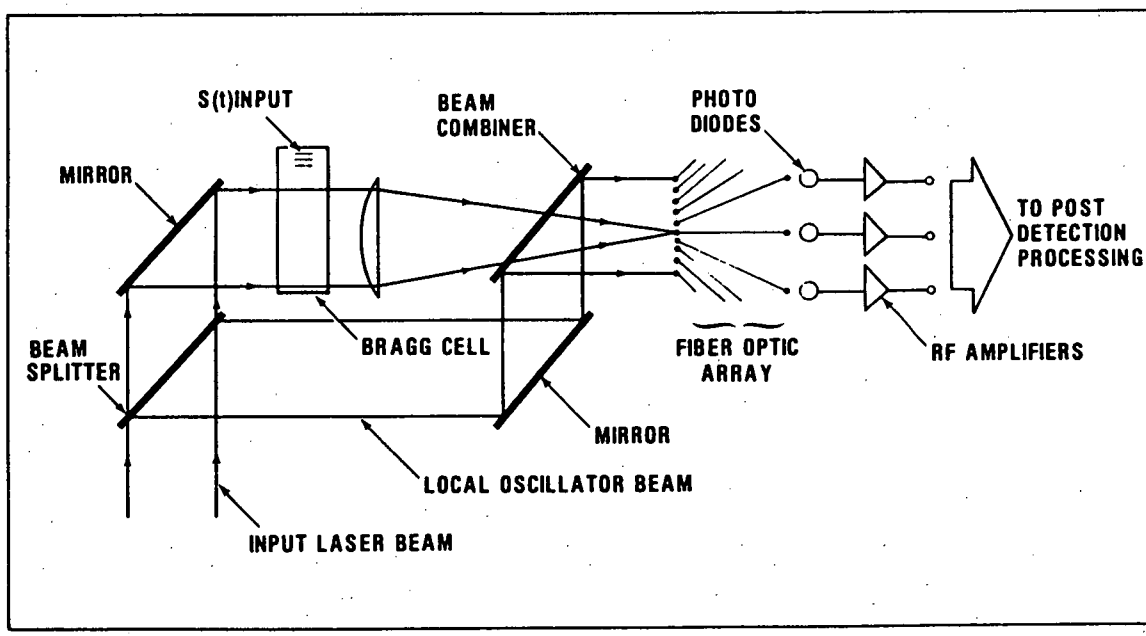


Fig.2-5: Coherent acousto-optical spectrograph. (After Berg et al., 1982).

The charge is proportional to the integrated incident light intensity so the detector performs both a square-law detection and integration function.

The mode of operation described corresponds to small signal operation of the Bragg cell. A signal which diffracts ten percent of incident light implies a linear dynamic range of 33 dB for the cell; one percent diffraction gives a 53 dB dynamic range (Hecht, 1977).

It can be shown (Milne and Cole, 1979) that the frequency resolution of an AOS is $\Delta f = 1/T_a$ where T_a, the aperture time, is the time taken for an acoustic wave to traverse the Bragg cell. Since T_a represents the duration of an input sample, the result is predictable. The normalized uncertainty of a spectral estimate, assuming the photodetector output fluctuations to be due solely to spectrometer input noise, is given by the radiometer equation

$$\sigma = \frac{1}{\sqrt{\Delta f \Delta t}} \quad ,$$

where Δt is the photodetector integration time.

In practical instruments there is a lower bound on input level below which sensitivity is degraded due to domination of detector noise. In addition, the photodetector array is required to have a wide dynamic range or the d.c. built up during integration of noisy signals will saturate the array before integration is complete (Turpin, 1981). At the high readout rates required for high time resolution spectrometry, noise due to the process of reading the CCD array becomes more significant and must be added to the inherent thermal and shot noise of the array. In astronomy applications the photodetector integration time is often a compromise between detector noise and saturation effects. Kellman et al. (1981) have constructed a detailed theoretical model of the processes involved in integrating spectra and have applied the model to various types of input signal.

Although the parameters of modern Bragg cells are quite suitable for radioastronomy (Cole and Milne, 1977), the AOS has several practical limitations. These can be summarised as:

(a) critical optical alignment with drift due to thermal and mechanical influences usually being a problem;

(b) non-uniform passband response due to the $(\sin x)/x$ form of the diffraction amplitude envelope (see, for example, Longhurst, 1973);

(c) limited dynamic range due mainly to modulator and detector characteristics (figures of 20 dB or less are reported).

In the form just described the AOS is a power spectrum analyser because the CCD array is a light intensity (square-law) detector. It is possible in principle to construct a more complex AOS with the capability of producing coherent spectra. Figure 2-5 shows the essential details of such an instrument. The device makes use of the shift in frequency of light in the diffracted beam relative to the incident beam frequency. In terms of a diffraction grating model, the frequency shift is due to the Doppler effect arising because of the motion of the grating in the Bragg cell. By using an optical heterodyne detection scheme, phase information can be recovered and the dynamic range can be considerably greater than for the basic AOS with intensity detection (Berg et al., 1982). The output of this form of AOS is directly analogous to the pre-detection output of a filterbank and, if the instrument is used in a polarimetry or other coherent processing role, the same considerations relating to the number of RF detectors and correlators apply.

With current technology the problems associated with a coherent AOS are acute. Alignment and stability problems are formidable and coherent optical detection is a technology in its infancy. Even if the problems with the spectrum analysers were overcome, the post-transformation hardware requirements are at least as great as with the filterbank approach. Future developments, such as integrated optical assemblies using surface acoustic wave rather than bulk wave diffraction gratings, may make the AOS approach viable in the next generation of coherent spectrum analysers.

2.4 DIGITAL SPECTRUM ANALYSERS

The most widely used method of digital spectrum analysis is based on the algorithm of Cooley and Tukey (1965) for the evaluation

of the discrete Fourier transform

$$F_k = \sum_{a=0}^{N-1} f_a \exp(-j2\pi ak/N) \qquad (0 \leq k \leq N-1) \ , \qquad 2.4.1$$

where f_a, the value of the a^{th} time sample of a continuous function f(t), is obtained in accordance with the usual restrictions applying to the sampling of bandlimited signals (Bracewell, 1978). The algorithm, known as the fast Fourier transform (FFT), gains speed by factoring N (the number of points or frequency resolution cells in the transform) into a product of integers,

$$N = r_1 r_2 \cdots r_m ,$$

corresponding to the breaking down of the one-dimensional signal x(t) into an m-dimensional form. If $N = r^m$ the transformation can be executed by performing a succession of r-point discrete Fourier transforms, the overall operation then being known as an radix-r FFT. In the simple case of r=2 the FFT algorithm restricts the number of multiplications needed to evaluate equation 2.4.1 to $2N.\log_2 N$, in contrast to the $4N^2$ or so necessary for direct computation of the transform.

In a radix-r FFT the number of elementary computation steps (or butterflies) needed to produce a complete transform is

$$C_r = \frac{N}{r} \log_r N \ .$$

As r increases, C_r decreases but the complexity of a butterfly increases. There is no single optimum value of r but r=2, 4, or 8 are common choices (Gold and Bially, 1973).

Notwithstanding the saving in computation time, hardware implementation of the algorithm is necessary for real-time analysis of bandwidths exceeding a few kilohertz. Most commercial real-time spectrum analysers based on hardware FFT processors are audio frequency instruments but high speed analysers are feasible. The FFT produces a complex spectrum so analysers using the algorithm can be used as the basis of a coherent processor.

Taking a radix-2 hardware processor as a reference, $(N/2).\log_2 N$ butterflies (each with an execution time of T_b) must be completed in the required computation time, T_c. For real-time operation T_c can be no greater than N/B, the time taken to fill the spectrum analyser input buffer when a sampling frequency of B is used. If complex sampling is employed (sin and cos input channels) B is also the maximum analysis bandwidth. Since

$$T_b = \frac{T_c}{\frac{N}{2}\log_2 N} ,$$

the butterfly time can be written as

$$T_b = \frac{2}{B \log_2 N} .$$

For a fixed butterfly time the product $B.\log_2 N$ is constant, so the designer must trade analysis bandwidth against frequency resolution.

Each butterfly is composed of at least one complex multiplication and two complex additions. A sub-system which performs a single butterfly is called a computational element. Fast bipolar multiplier IC's can form a 16-bit product in around 150 ns and bipolar or ECL memory access times are of the order of 10 ns. Addition times are also around 10 ns. With T_b ~350 ns and N=1024, B=570 kHz. For N=64, B~950 kHz. If B=30 MHz and N=64, T_b needs to be about 11 ns. Clearly, serial processing is not practical in wideband instruments and some form of parallel processing must be used.

A practical method of incorporating parallel architecture is to implement a pipeline (or cascade) processor in which $\log_2 N$ butterflies are computed in parallel. Hence, each computational element computes only N/2 butterflies. A further relaxation in T_b can be obtained by constructing a fully parallel processor in which all N/2 butterflies in a single stage of the FFT are computed simultaneously (Johnston, 1983). The ultimate in parallelism is an array processor which combines parallel and pipeline principles to produce an instrument able to calculate all $(N/2).\log N$ butterflies in parallel, giving $T_b=T_c$. Such a system requires extensive hardware and is not often employed.

The survey of digital Fourier transformers conducted by Roberts et al. (1980) indicates that, with one exception, the use of special hardware elements and the placement of few restrictions on power consumption, cost or physical size has not resulted in 1K-point analysers with a computation time of less than 1 ms, corresponding to a real-time bandwidth limit of 1 MHz with complex sampling.

In a digital processing system making use of the inherently coherent nature of the analysis (such as a polarimeter) two continuously operating channels are necessary. The amount of hardware is daunting, especially when the need for post-transformation processing is taken into account.

The FFT approach becomes more manageable when implemented using VLSI technology. Although common VLSI processes are still far too slow for direct realization of a high time resolution analyser, the use of bipolar processes and highly parallel architecture can result in efficient, wideband instruments (McTaggart, 1982). In addition, the multiple project chip (MPC) concept is gaining popularity so it is likely that specialised integrated circuits (such as those needed for post-transformation processing in a polarimeter) will become viable, even in situations where production line outputs are small.

The potential of gallium arsenide (GaAs) LSI technology should also be borne in mind. A recent prototype GaAs IC forms the 16-bit product of two 8-bit numbers in 5.25 ns (an equivalent propagation delay of 150 ps per gate) and by implementing a lookahead scheme together with Booth's multiplication algorithm, the same device could produce a 16-bit product in 2.25 ns (Lee et al., 1982). High time resolution FFT coherent processors are therefore certain to become practical and will enable designers to take advantage of the high accuracy and wide dynamic range achievable with digital signal processing methods.

As an alternative to direct computation, digital correlators can be used to allow indirect estimation of spectra. Correlation spectrum analysers are widely used in radioastronomy where the power spectrum of noisy, wideband signals is often required (Goldstein, 1962; Weinreb, 1963). The correlation approach is based on the Wiener-Khintchine theorem which states that, for both deterministic and stochastic

signals, x(t), a Fourier transform relationship holds between the signal power spectrum, P(f), and the signal auto-correlation function, $R(\tau)$. That is,

$$P(f) = \int_{-\infty}^{\infty} R(\tau) \exp(-j2\pi f\tau)\, d\tau \quad . \qquad 2.4.2$$

Since practical measurements cannot involve infinite limits in τ, the estimate is subject to error. In practice, the expression

$$R(\tau) = \langle x(t)\, x(t-\tau) \rangle \qquad 2.4.3$$

(where <> denotes a time average) can be estimated and used as the basis for the Fourier transformation process. In this case $R(\tau)$ is the expectation value of the product of two x(t) values displaced in time by τ and can be deduced for a large range of τ (Roberts et al., 1980).

If the signal x(t) is noiselike in the sense that the joint probability density function of the two multiplicands in equation 2.4.3 is bivariate Gaussian, Van Vleck and Middleton (1966) show that

$$R(\tau) = \sin[\tfrac{\pi}{2} R'(\tau)]$$

where $R'(\tau)$ is the auto-correlation function of the signal derived by clipping or hard-limiting x(t). Since many signals in radioastronomy exhibit very low signal-to-noise ratio the assumption of Gaussian statistics is often well justified and, if the signal is clipped or quantized to the one-bit level, very simple logic circuits can be used to deduce $R'(\tau)$ from which $R(\tau)$ can be recovered. Fourier transformation then yields P(f).

The effect of using one-bit correlation as a means of obtaining spectra for cosmic emission having finite and variable signal-to-noise ratio is difficult to quantify. Cordes et al. (1983) have examined the effect of one-bit analysis on pulsar signals which are modelled as amplitude modulated white noise (Rickett, 1975). They conclude that one-bit correlation leads to a distortion of the true spectrum with the shape of the computed spectrum depending on the signal-to-noise ratio.

The distortion is largely removable if millisecond estimates of $R'(\tau)$ are recorded (or, presumably, if high resolution spectra are produced in real time) and estimates of on-pulse and off-pulse intensity are made. No equivalent study for transient solar and Jovian emission is available although Grognard and McLean (1973) used a one-bit correlation approach in a search for linear polarization in type III solar bursts.

A basic auto-correlation spectrometer is shown in figure 2-6. The band-limited signal, $x(t)$, has zero mean and is one-bit quantized at a frequency at least equal to the Nyquist frequency. The sampler output, representing the sign of $x(t)$ at the sampling instant, is fed into the shift register. After each clock cycle, multiplier M_n forms the product of the current sample and a delayed sample to give the product $x'(t).x'(t-n\Delta t)$ where Δt is the clock period. Since x' is a one-bit variable the multipliers are simple coincidence detectors, usually exclusive-OR or exclusive-NOR logic gates. The result of the multiplication (either logical 0 or 1) is added to the total in the counter, C_n. It is only necessary to count either logical-1 or logical-0 results, each count being normalized by C_0, the total number of products for a delay $n\Delta t$. If the counters accumulate coincidences,

$$R'(n\Delta t) = 2\left(\frac{C_n}{C_o}\right) - 1$$

(Cooper, 1976).

The values of $R'(n\Delta t)$ are read into either a computer or a hardware processor, the continuous correlation function $R(\tau)$ recovered using the Van Vleck relationship, and the final Fourier transformation performed to yield the power spectrum, $P(f)$. Since integration in the correlator can be essentially continuous, slow Fourier transformation leads only to a loss in time resolution, not a reduction in sensitivity as is the case with a slow FFT analyser.

For m counting channels the frequency resolution of the correlation spectrometer is given approximately by

$$\Delta f = \frac{1}{m\Delta t}$$

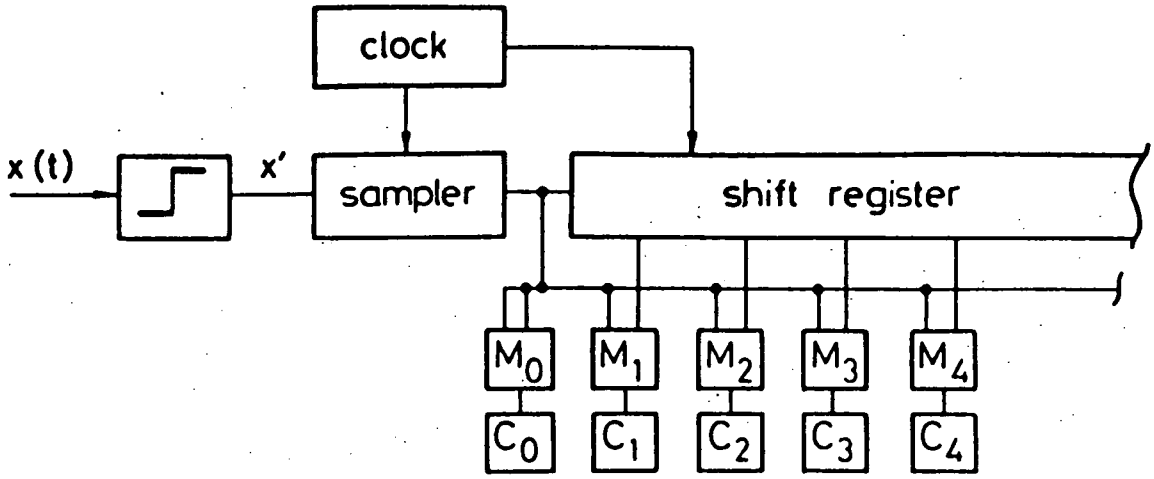


Fig.2-6: Basic one-bit autocorrelation spectrometer. (After Roberts et al., 1980).

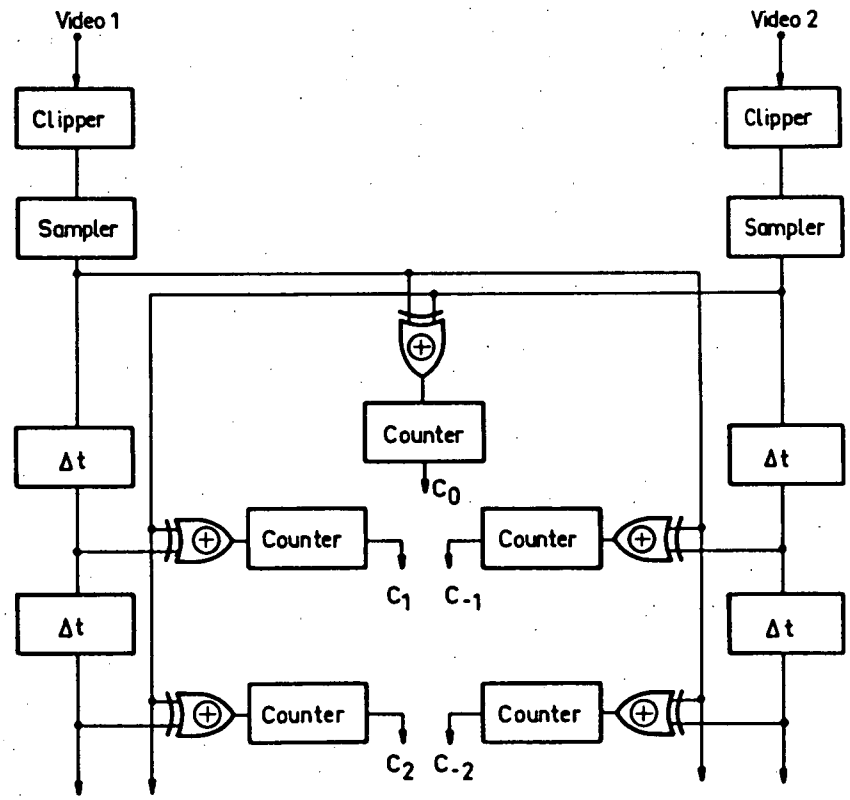


Fig.2-7: One-bit cross-correlator. The EXCLUSIVE-OR gates are multiplier elements and the delay chains can be shift registers as shown in figure 2-6. (After Cooper, 1976).

and the normalized uncertainty of the spectral estimate is approximately

$$\sigma = \frac{\beta}{\sqrt{T\Delta f}}$$

if the signal statistics are stationary over the integration time T. The degradation factor, β, relative to a filterbank spectrometer arises because of the information loss involved in the one-bit quantization process. Van Vleck and Middleton (1966) show that $\beta = \pi/2$ or 1.6 for a one-bit correlator. Although quantization to more levels is possible, the correlator hardware is complicated considerably and the gain may not be worthwhile. For example, two bit quantization gives β=1.1.

The previous discussion relates directly to auto-correlation spectrometers which produce power spectra and which are therefore unsuitable as the sole basis of a coherent processor. From the filterbank spectrometer arrangement (figure 2-3) it is clear that a correlation polarimeter forms two cross-spectra representing the Stokes parameters U and V when the instrument is connected to orthogonal antennas sensitive to linear polarization. Digital correlation methods can also be used to derive cross-spectra. In this case $R(\tau)$ is the cross-correlation function of two separate signals, x(t) and y(t). The expression analogous to equation 2.4.3 is

$$R(\tau) = \langle x(t)\, y(t-\tau) \rangle$$

and $R'(\tau)$, generated by one-bit quantizing x(t) and y(t), is used to derive $R(\tau)$. The Fourier transform of $R(\tau)$ then gives the cross-spectrum of the signals x(t) and y(t).

In general, $R(\tau)$ is an asymmetrical function of τ so the cross-spectrum is complex. Cooper (1976) illustrates the convenience of generating the function $R(-\tau)$ as well as $R(\tau)$, allowing the cross-correlation function to be separated into its even and odd components. The Fourier cosine transform of the even function then yields the spectrum of the Stokes parameter U and the sine transform of the odd component gives the V spectrum.

Figure 2-7 shows a practical digital cross-correlator. Two shift register delay paths are used to allow coincidence detection for

both positive and negative delays. As with an autocorrelation spectrometer, the observatory computer or a dedicated processor forms the continuous correlation functions and performs the required Fourier transformations.

For a polarimeter or other coherent analysis system three spectrum analysers are needed: two autocorrelation (intensity) analysers and a cross-correlation analyser. In addition, a processor is required to acquire data from each correlator, recover the continuous correlation functions and perform the Fourier transforms.

Even if the correlator hardware is designed to handle sampling frequencies of greater than, say, 50 MHz (needed for wide-band operation), the Fourier transformation process produces a data "bottleneck" at the correlator output. At 1 ms time resolution (barely adequate for many purposes) the Fourier transformer is required to produce a spectrum at least every 330 μs, assuming one transformer is used. A typical general purpose computer with an array processor attachment takes around 5 ms to produce a 1K-point spectrum, so is an order of magnitude too slow even if data transfer and acquisition time is neglected.

Dedicated hardware FFT processors have been discussed earlier and based on the results quoted, the speed specification for the Fourier transformation may be almost met if very advanced hardware is used. As with FFT instruments, the correlator approach will become more viable with the availability of MPC VLSI programs although it is probable that correlation spectrometry will be eventually supplanted by hardware FFT methods in many applications.

At present, digital correlation spectrum analysers are highly effective in most areas of radioastronomy, including those which require coherent processing to be performed. For high time resolution, real-time analysis of wide bandwidths, speed limitations make the approach impractical with current technology.

2.5 THE SCANNING SPECTRUM ANALYSER

The scanning spectrum analyser is the standard instrument for spectral analysis in many RF test laboratories. The analyser is essentially a superheterodyne receiver, the local oscillator (LO)

frequency of which is varied (usually linearly) with time. The basic configuration is illustrated in figure 2-8.

As the LO frequency is changed the detected receiver output is displayed on a CRT vertical axis. If a voltage proportional to LO frequency is applied to the horizontal axis, the display shows spectral amplitude (or power if a square-law detector is used) as a function of frequency. The simplistic assumption that the resolved bandwidth, Δf, not be traversed in a time less than about $1/\Delta f$ leads to the approximate relationships

$$\Delta f \sim \sqrt{k} \qquad 2.5.1$$

and

$$B \sim \Delta f \sim \sqrt{k} \quad , \qquad 2.5.2$$

where B is the optimum IF bandwidth (in Hertz) and k is the scan rate in Hertz per second.

Butson et al. (1966) show that if the phase response of the IF filter is linear, the above relationships are given more accurately by

$$\Delta f = 0.7\sqrt{k}$$

and

$$B = 1.6\sqrt{k}$$

if the IF passband is rectangular.

The equations relating IF bandwidth, frequency resolution and scan rate are valid when the signal characteristics are stationary, at least over a time-scale greater than the scan time T. For short duration signals the scanning analyser is not satisfactory since it is possible for such signals to appear and disappear without registering at the analyser output. This low probability of intercept (POI) is inherent and if the scan rate is increased in an attempt to raise the POI, the frequency resolution (given by equation 2.5.1) is degraded. The need to trade probability of intercept for frequency resolution is a fundamental disadvantage of scanning spectrum analysers in applications where short duration signals are encountered, such as in the observation of dynamic cosmic radio emission.

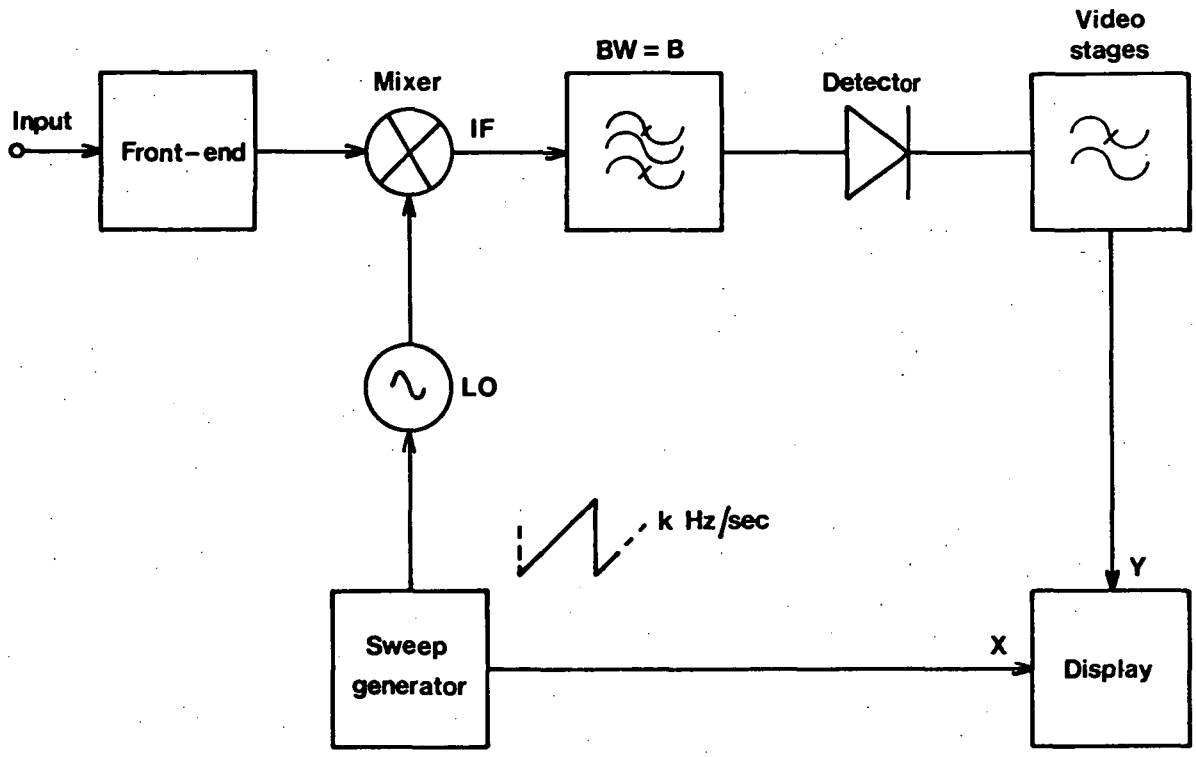


Fig.2-8: Conventional scanning spectrum analyser.

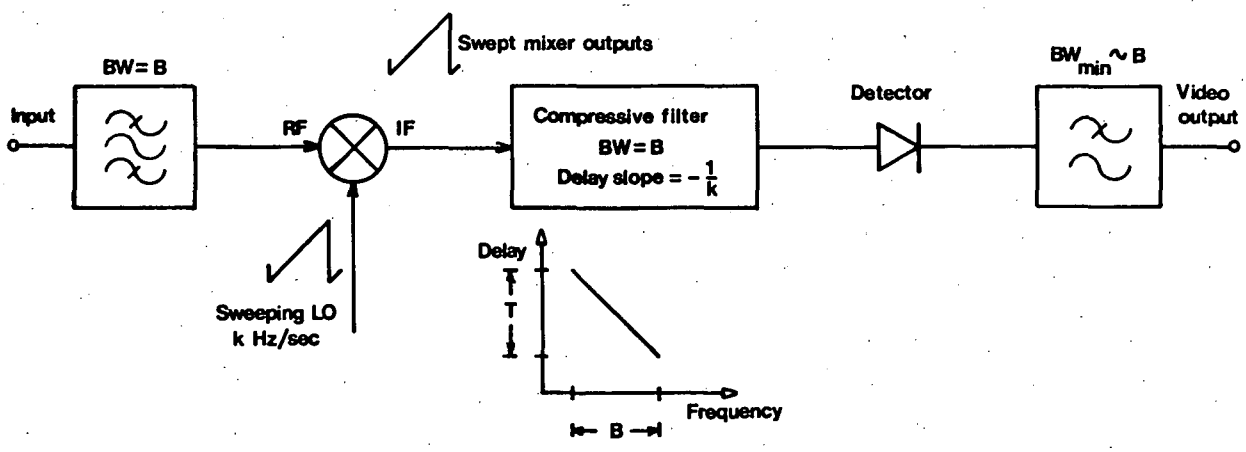


Fig.2-9: Compressive receiver spectrum analyser. The input bandpass filter is not necessary in all cases but is included to simplify the preliminary discussion in this chapter.

A second major disadvantage of the scanning analyser is its waste of signal energy: if J frequency cells are resolved the time spent observing any one cell is only 1/J of the total observation time. Although it is possible to envisage a coherent processor or polarimeter based on two scanning analysers, the disadvantages mentioned make any such system unattractive in all but a few specialised roles.

2.6 PULSE COMPRESSION SPECTRUM ANALYSIS

In this section the development of the pulse compression or compressive receiver spectrum analyser from the scanning analyser is illustrated and its advantages outlined. The description is meant as an introduction since the theoretical basis of compression analysers is discussed in chapter four.

From equation 2.5.2, the IF bandwidth of a scanning spectrum analyser should be approximately equal to the square root of the scan rate if optimum frequency resolution is to be obtained. This conclusion rests partially on the assumption that the IF phase response of the analyser is linear. It has become apparent (Butson et al., 1966; Harrington and Nelson, 1974) that better frequency resolution can be obtained by allowing the IF phase response to deviate from a linear characteristic. The essence of the pulse compression technique is the introduction of a differential time delay between frequency components of the IF band emerging from the mixer of a scanning spectrum analyser. If this frequency dependent delay compensates exactly for the dispersion introduced by the scanning LO, all the signal energy tends to emerge simultaneously from the compressive IF filter.

The basic compressive receiver is shown in figure 2-9 and the compression operation for CW signals falling within the receiver processing bandwidth is illustrated in figure 2-10. It is apparent that, for a linear LO scan with slope k Hertz per second, the IF compressive filter must provide a time delay which is a linear function of frequency. If mixing a CW input with the sweeping local oscillator causes the signal to remain in the IF passband for time T, the IF filter group delay slope must be -1/k second per Hertz if the total delay across the passband is to be T and optimum compression is to occur.

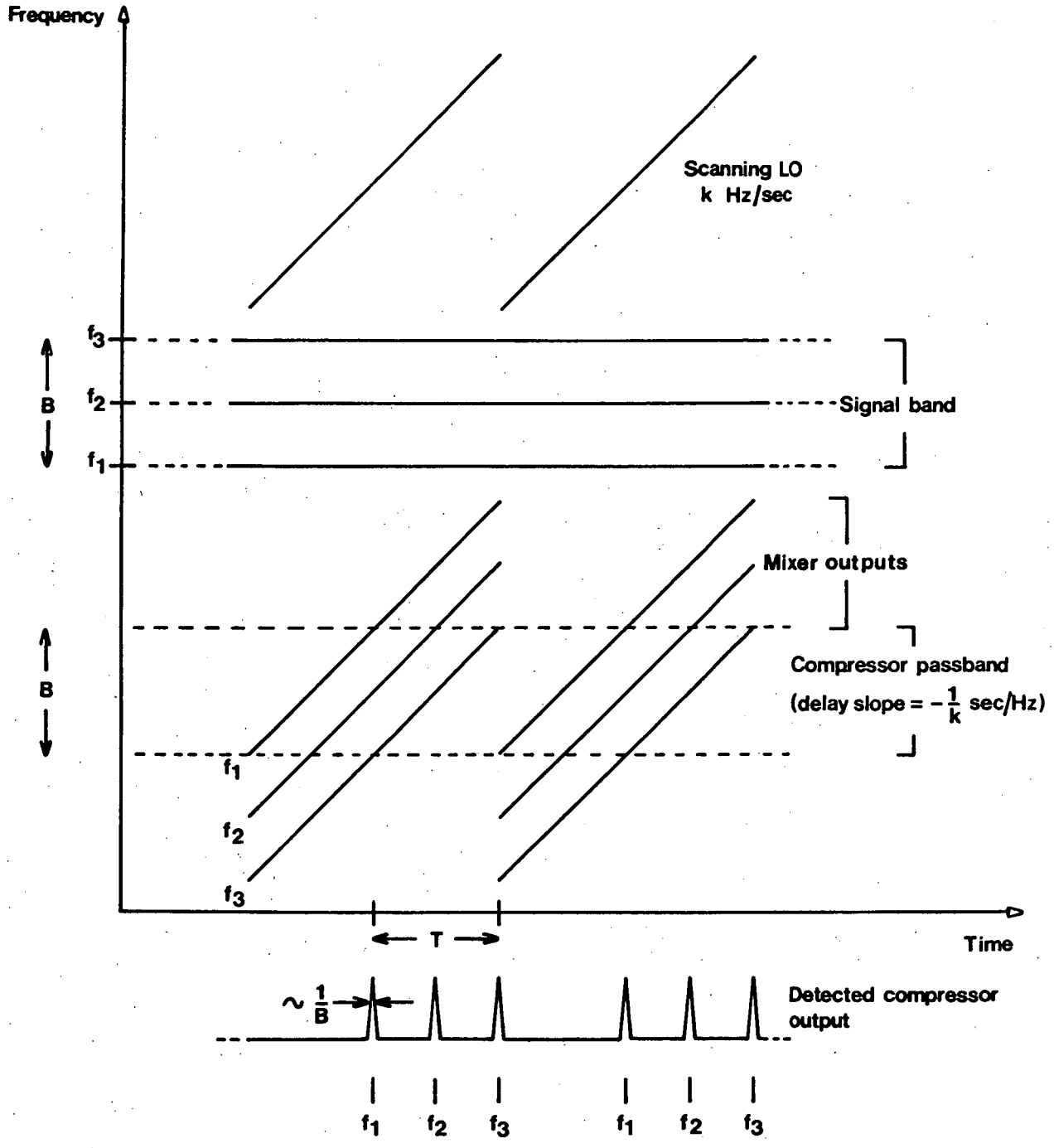


Fig.2-10: Frequency-time diagram showing the operation of a compressive receiver.

Phase is proportional to the integral with respect to frequency of group delay, so the linear delay requirement corresponds to a quadratic phase-frequency specification for the compressive filter. Further, the derivation given in section 3.2 will show that the compressive filter impulse response has the form of the filter's input signal reversed in time, leading to the view that pulse compression is a particular example of matched filtering.

Separation in frequency at the compressive receiver input leads to separation in time of emergence of compressed pulses at the output (figure 2-10), so the device functions as a spectrum analyser. The spectrum emerges as a time waveform occupying a time interval T. For a CW input the compressed pulse is about 1/B seconds wide in time, giving T/(1/B) or TB individually discernible frequency resolution cells. The product TB is called the time-bandwidth product of the spectrum analyser.

The compression technique gives an IF filter output pulse which is substantially narrower than B/k, the time taken for the signal to be swept across the IF passband in a normal scanning analyser. The pulse width is of the order of 1/B so, in principle, the frequency resolution can be increased indefinitely by increasing the bandwidth of the dispersion (or expansion) and compression process. However, the reciprocal spreading relationship ($\Delta f \Delta t \sim 1$) still applies and an increase in frequency resolution is accompanied by a corresponding decrease in time resolution. The important point is that the process eliminates the dependence of the output pulse width on the slope of the LO signal, making it possible to achieve frequency and time resolutions set only by the spreading relationship. This means that it is possible to sweep through the bandwidth to be analysed many times faster than is possible with a conventional scanning spectrum analyser of equivalent frequency resolution (see section 4.4 for a comparison). The compressive receiver probability of intercept is therefore much higher than that of a scanning analyser and this attribute has made the instrument popular with designers of military radar intercept or surveillance systems.

Signal energy must be conserved so forcing all signal energy to emerge from the IF filter simultaneously increases the peak signal amplitude at the output. Noise components of the IF signal are uncorrelated with respect to frequency and the differential delay has no effect on the

mean output noise level. Thus, although the dispersive IF filter is a passive, linear device and therefore cannot increase the signal-to-noise power ratio, it does increase the ratio of the peak signal to mean noise amplitude for CW or quasi-CW signals. This effect serves to enhance signal detectability in many applications. A quasi-CW signal is a signal which is present for a duration of the order of the LO sweep time and the frequency of which remains substantially constant over the same time scale.

Adjusting the operational cycle (figure 2-10) of the analyser so that there is no dead time between successive output spectra ensures that all the energy of a continuous input signal contributes to the analyser output. In sensitivity terms, a compressive receiver therefore resembles more closely a filterbank analyser than a scanning type since, in principle, no signal energy need be wasted in order to obtain given bandwidth and resolution parameters.

A further advantage of the compressive receiver, little exploited to date, is the coherent nature of the spectrum analysis performed. It is shown in chapter four that, for the analyser illustrated in figure 2-9, the pre-detection output consists of the Fourier transform of the input multiplied by a high frequency component at the IF filter centre frequency. This RF term, the phase of which is related to the phase of signals at the analyser input, can be used as a reference for processing in a correlation polarimeter spectrometer or other coherent processor. It is also possible to construct a two-channel adding polarimeter using compressive receivers because the coherent spectrum analysis process allows an IF phasing unit (figure 1-4) to be used.

Figure 2-11 shows a basic correlation polarimeter using two compressive receiver spectrum analysers. The outputs of the two analysers, which are functional equivalents of the filterbanks in figure 2-3, are multiplied directly to give in-phase and quadrature products having video (baseband) components proportional to the product Stokes parameters. Since all four outputs emerge simultaneously in serial analog form, time sampling in each processor output channel can be used to acquire data from successive frequency resolution cells in the four spectra.

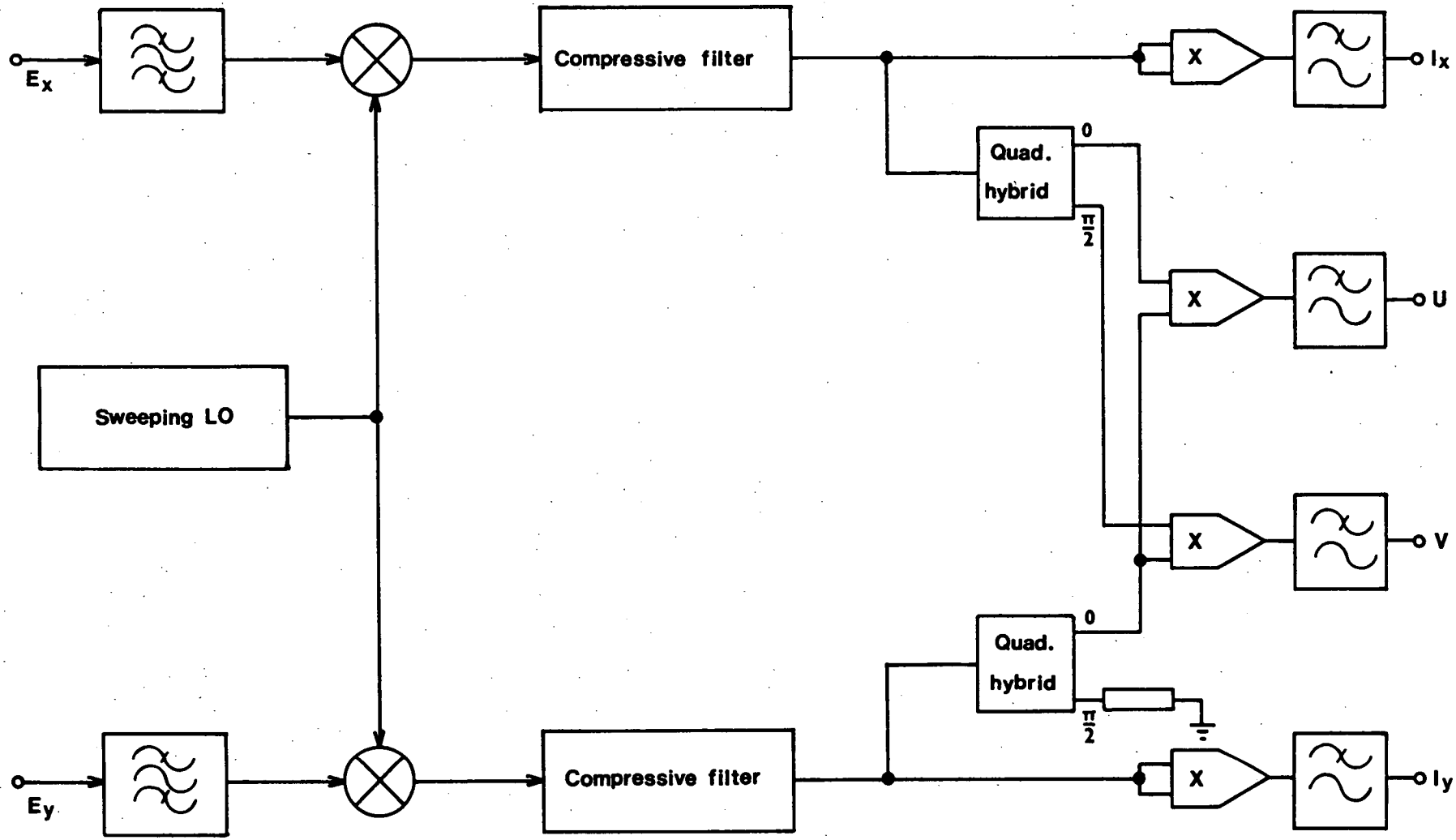


Fig.2-11: Correlation Stokes polarimeter spectrometer using two compressive receivers. Note the similarity to the narrowband polarimeter shown in figure 1-2.

The coherent processor outlined is the basis of the correlation polarimeter described in chapter seven. It offers several advantages over other approaches reviewed in earlier parts of this chapter but, at the same time, construction of any compressive receiver spectrometer able to be interfaced to a general purpose computer is a novel and difficult engineering undertaking. Before proceeding to a formal statement of the signal processing techniques invoked and a detailed description of the instrumentation, it is worth summarising the advantages of the new device and outlining the major problems in implementing the design sketched in figure 2-11. Many characteristics are common to a system using compressive receivers as power spectrum analysers.

The advantages, for a processor constructed around surface acoustic wave (SAW) dispersive filters, are:

(a) the ability to achieve sensitive,high time resolution analysis over a wide bandwidth without trading more frequency resolution than required by the reciprocal spreading relationship (time-bandwidth products of up to 10 000 are feasible);

(b) low spectrum analyser hardware complexity since high quality, reproducible, SAW dispersive filters are commercially available at prices which give a total system price equal to or less than that of alternative analysers;

(c) low post-transformation processing complexity because the serial nature of each spectrum analyser output means that only one set of detectors and correlators is needed;

(d) high dynamic range (25 dB is routinely achievable);

(e) simple alignment;

(f) stable performance with respect to thermal and mechanical influences, meaning that no adjustment is required after initial alignment.

The major engineering demands associated with a processor based on compressive receivers centre on development of the following items:

(a) accurate, four-quadrant, bi-linear (equal gain from both inputs) analog multipliers (for use as correlators or detectors) having input and output bandwidths wide enough to accommodate the processing bandwidth of compressive receivers;

(b) high speed digital equipment (with sampling and processing clock frequencies of the same order as the receiver bandwidth) capable of acting as a dedicated intermediate data processor between the fast SAW instrument and ordinary data acquisition computers;

(c) a high speed data transfer system allowing the intermediate processor to communicate with a data acquisition computer.

As well as hardware development, an operational protocol must be devised to allow interconnection of three signal processing stages (SAW coherent processor, intermediate digital processor and computer) to form an effective real-time analysis system.

These problems are major ones and have not been surmounted (at least in an economically realistic manner) even given several military and associated industry research programs. This thesis outlines a first attempt at a coherent signal processor based on SAW compressive receivers and capable of being interfaced to standard computers.

The pulse compression operation is the basis of compressive receiver spectrum analysis and is examined in detail in the next chapter.

CHAPTER 2 - REFERENCES

Berg, N.J., Casseday, M.W., and Abramovitz, I.J.: Acousto-optic processing increases E.W. capabilities. Microwave Systems News, Vol.12, No.2, Feb. 1982, pp.106-118.

Bergland, GF.D.: Fast Fourier transform hardware implementations - an overview. IEEE Trans. Audio Electroacoust., Vol.AU-17, No.2, June 1969, pp.104-108.

Bracewell, R.: "The Fourier transform and its applications" Ch.10, McGraw-Hill, 1978.

Butson, P.C., Lucas, W.J. and Thompson, G.T.: Theoretical assessment of the use of pulse compression in a panoramic receiver. Proc. IEE, Vol.113, No.5, May 1966, pp.725-739.

Cole, T.W. and Milne, D.K.: An acousto-optical radio spectrograph for spectral integration. Proc. Astron. Soc. Aust., Vol.3, No.2, Sept. 1977, pp.108-111.

Cooley, T.W. and Tukey, J.W.: An algorithm for the machine computation of complex Fourier series. Math. of Comp., Vol.19, No.90, 1965, pp.297-301.

Cooper, B.F.C.: Autocorrelation spectrometers. In "Methods of experimental physics" (M.L. Meeks ed.). Vol.12, Ch.3.5, Academic Press, 1976.

Cordes, J.M., Weisberg, J.M. and Boriakoff, V.: An attempt to resolve pulsar magnetospheres using interstellar scintillation. Astrophys. J., Vol.268, 1983, pp.370-380.

Ellis, G.R.A.: Private communication, 1984.

Gold, B. and Bially, T.: Parallelism in fast Fourier transform hardware. IEEE Trans. Audio Electroacous., Vol.AU-21, No.1, Feb. 1973, pp.5-16.

Goldstein, R.M.: A technique for the measurement of the power spectra of very weak signals. IRE Trans. Space Elec. Telemetry, Vol.SET-8, No.2, June 1962, pp.170-173.

Grognard, R.J.M. and McLean, D.J.: Non-existence of linear polarization in type III solar bursts at 80 MHz. Solar Phys., Vol.29, 1973, pp.149-161.

Harrington, J.B. and Nelson, R.B.: Compressive receiver uses SAW devices. Microwave Journal, Vol.17, No.9, Sept. 1974, pp.57-62.

Hecht, D.L.: Multifrequency acousto-optic diffraction. IEEE Trans. Son. Ultrason., Vol.SU-24, No.1, Jan. 1977, pp.7-18.

Johnston, J.A.: Parallel pipeline fast Fourier transformer. Proc. IEE, Vol.130, Pt.F, No.6, Oct. 1983, pp.564-569.

Kellman, P., Shaver, H.N. and Murray, J.W.: Integrating acousto-optic channelized receivers. Proc. IEEE, Vol.69, No.1, Jan. 1981, pp.93-100.

Lee, F.S. et al.: A high speed LSI GaAs 8×8 bit parallel multiplier. IEEE J. Solid State Circ., Vol.SC-17, No.4, Aug. 1982, pp.638-646.

Longhurst, R.S.: "Geometrical and physical optics" Ch.12, Longman, 1973.

McTaggart, R.: High bandwidth FFT's. University of Edinburgh Report, 1982.

Milne, D.K. and Cole, T.W.: A broadband radio astronomy spectrometer. Proc. IREE, Vol.40, No.2, March 1979, pp.43-47.

Rickett, B.J.: Amplitude modulated noise: an empirical model for the radio emission received from pulsars. Astrophys. J., Vol.197, 1975, pp.185-191.

Roberts, J.B.G., Moule, G.L. and Parry, G.: Design and application of real-time spectrum analyser systems. Proc. IEE, Vol.127, Pt.F, No.2, April 1980, pp.76-91.

Stuart, R.D.: "An introduction to Fourier analysis", Ch.4, Methuen, 1961.

Tiuri, M.: Radio astronomy receivers. IEEE Trans. Ant. Prop., Vol.AP-12, No.7, Dec. 1964, pp.930-938.

Turpin, T.M.: Spectrum analysis using optic processing. Proc. IEEE, Vol.69, No.1, Jan. 1981, pp.79-92.

Van Vleck, J.W. and Middleton, D.: The spectrum of clipped noise. Proc. IEEE, Vol.54, No.1, 1966, pp.2-19.

Weinreb, S.: A digital spectral analysis technique and its application to radioastronomy. MIT Tech. Report, No.412, 1963.

CHAPTER 3 - PULSE COMPRESSION

3.1 INTRODUCTION

The pulse compression process was developed originally as a means of maintaining both sensitivity and range resolution in military radar systems and, even today, much of the research associated with compression hardware and techniques is related to radar. The essence of the radar problem is that target detectability is related directly to the amount of energy transmitted, which is in turn proportional to the duration of the transmitted pulse. However, the time, and therefore the range, resolution is directly proportional to the bandwidth occupied by the transmitted signal. If the pulse is extended to ensure good sensitivity the bandwidth ($B \sim 1/T$) is reduced and so the range resolution is degraded.

The basis of the pulse compression technique, first reported in detail by Klauder et al. (1960) and Cook (1960), is an artificial increase in the bandwidth occupied by a transmitted pulse. A pulse of given duration therefore yields higher range resolution than would otherwise be possible. The usual method of bandwidth extension is to frequency modulate the pulse (usually with a linear ramp), giving rise to a swept frequency pulse signal known as a "chirp". Depending on the sense of the frequency modulation an "up-chirp" or a "down-chirp" may be produced. In a radar system (figure I-10, appendix one) a compressive filter is used to construct a pseudo-impulse out of the chirp arriving at the receiver. The time between the appearance of successive compressed pulses is related to the separation of targets within the radar beam.

This chapter reviews theoretical aspects of pulse compression. The results derived are directly relevant to the spectrum analyser application discussed in detail in chapter four. Frequent referral is made to the original sources mentioned above although some results have been derived using approaches which are more direct than those originally employed. Newer literature is also cited, particularly in the discussion of matched filter chirp weighting.

3.2 THE PULSE COMPRESSION OPERATION

In this section the basic concepts of pulse compression are outlined. Consider the up-chirp signal

$$f(t) = w(t)\ \exp[j2\pi(f_0 t+\tfrac{1}{2}kt^2)] \quad , \qquad 3.2.1$$

where w(t) is an arbitrary envelope (or "weighting") function. The instantaneous frequency, defined as the rate of change of phase, is given by

$$f_i = \frac{1}{2\pi}\frac{d\phi}{dt} = f_0+kt \quad , \qquad 3.2.2$$

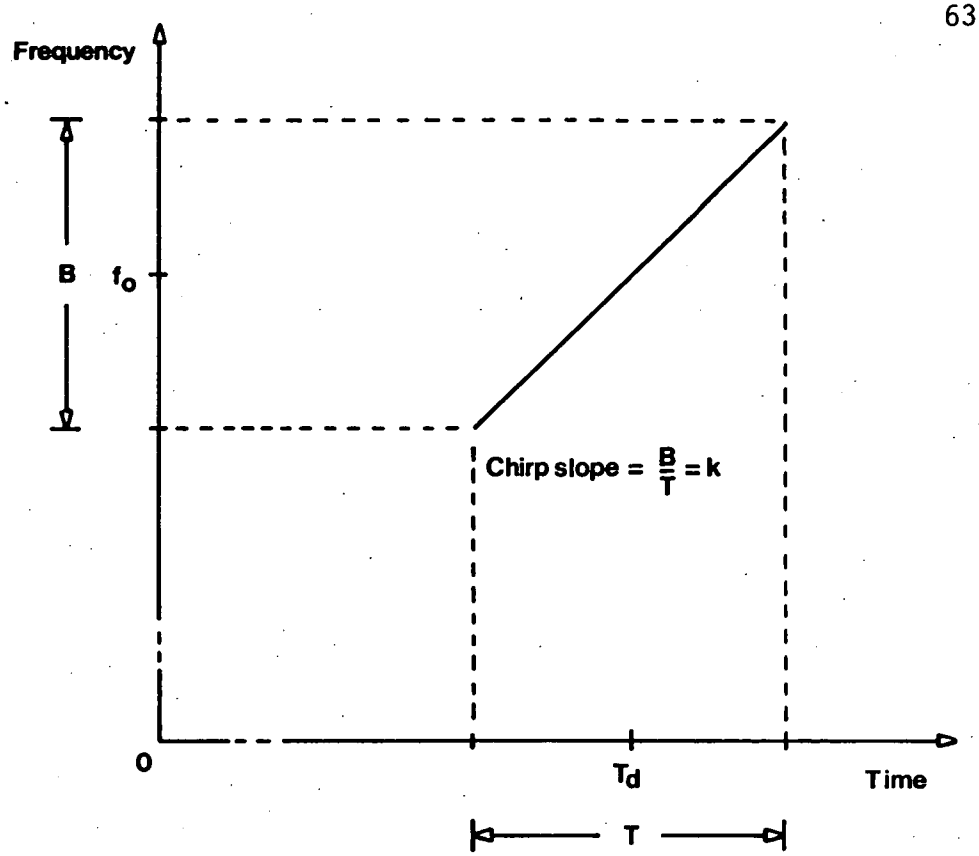
where k is the chirp slope in the cyclic frequency-time plane.

The signal is illustrated graphically in figure 3-1(a). In time T the frequency moves over a total swing B. The product TB is called the time-bandwidth product of the chirp and is an indicator of its signal processing potential and of the physical complexity of the device producing it. The absolute delay time, T_d, arises from realizability requirements of real networks used in chirp generation. In forthcoming analyses T_d is set to zero with no loss of generality in the results derived. Thus, the chirp signal is defined over the interval $-T/2 \leq t \leq T/2$. If w(t) is the rectangular gating function (figure 3-1(b)) defined by

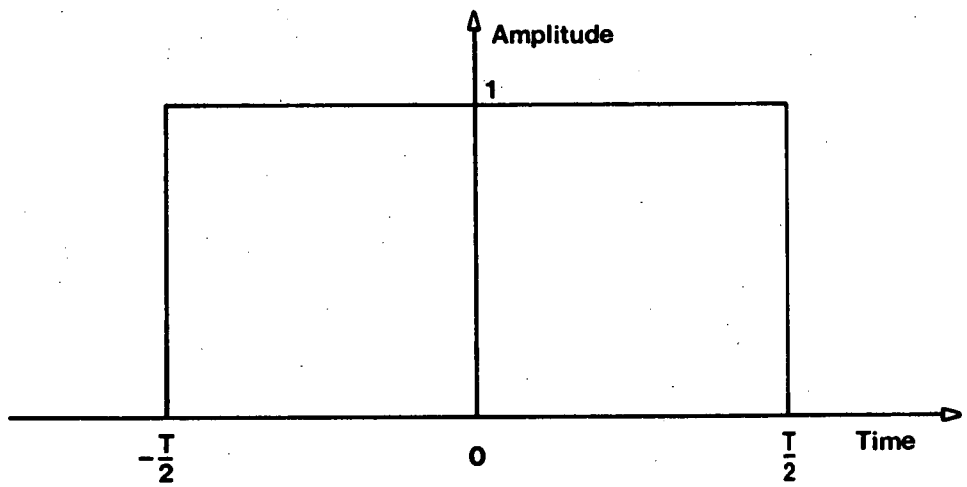
$$w(t) = \Pi\left(\frac{t-\frac{1}{2}T}{T}\right) \quad , \qquad 3.2.3$$

the chirp signal is said to be un-weighted or uniform.

Consider now the form of the network response required to compress or collapse the chirp defined by equation 3.2.1. It is intuitively clear that the use of a compression network having a group delay versus frequency characteristic which is the inverse of the dispersion characteristic of the chirp should produce a compressed output. In figure 3-2 the up-chirp input means that low frequencies at the start of the pulse enter the compressor first, but the frequency dependent delay characteristic of the compressor is such that these frequency components are delayed most. High frequency components entering the compressor at the end of the pulse experience least delay. The overall effect is



(a)



(b)

Fig.3-1: Graphical representation of a chirp signal. The frequency-time characteristic of an up-chirp is shown in (a). The envelope function for the un-weighted chirp is shown in (b). The absolute delay, T_d, is set to zero for analysis purposes.

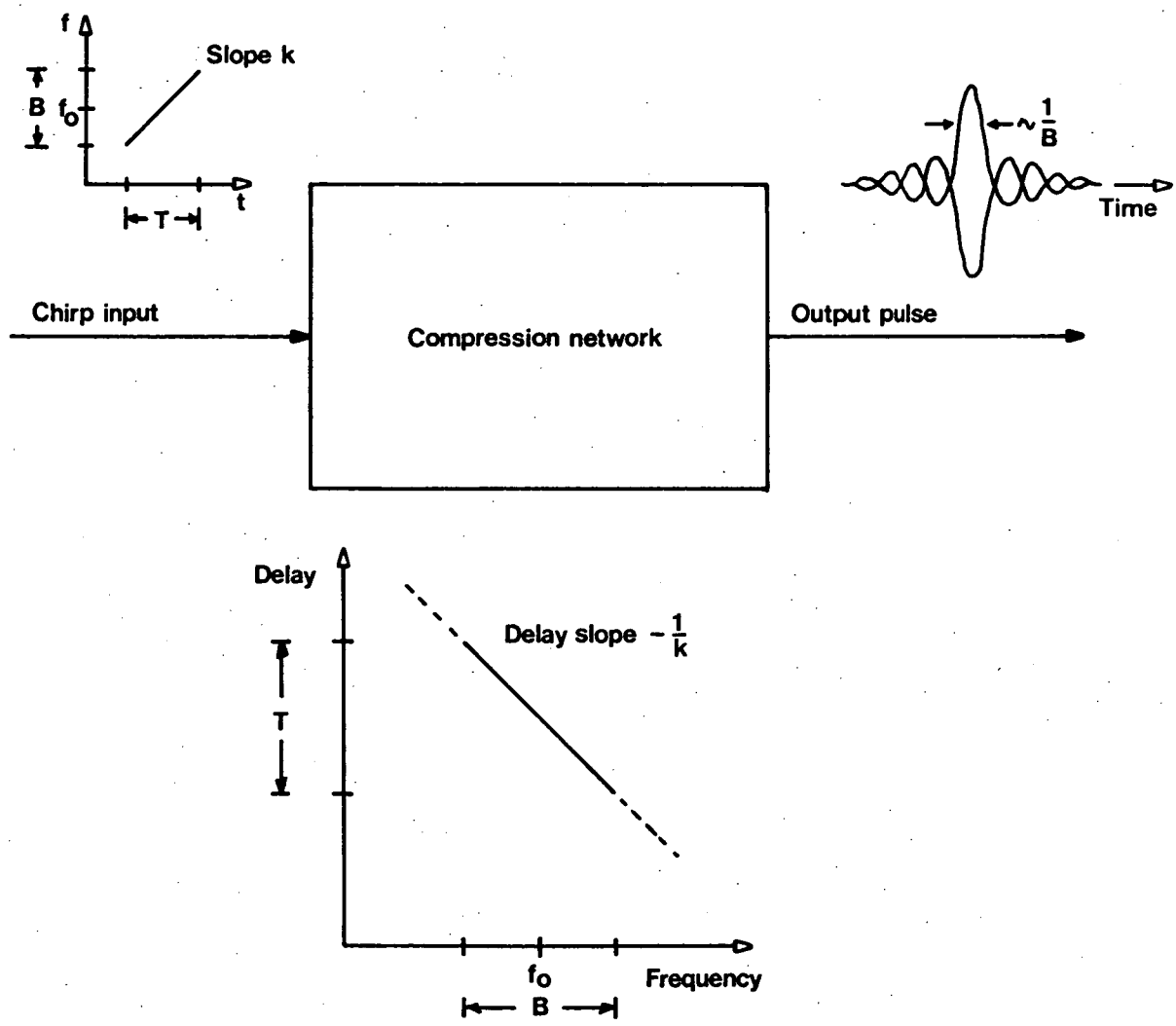


Fig.3-2: The pulse compression operation. Frequency dispersion in the input chirp is removed by the compressor, resulting in an output pulse width of the order of the inverse bandwidth of the compression process.

that dispersion present in the input signal is removed by the network. The expected compressor output is a pulse with a duration of about 1/B, the reciprocal chirp bandwidth. It is the exact form of this output which is now deduced.

Suppose the compression filter is an infinite bandwidth, lossless, phase equalizer with an admittance function

$$H(f) = \exp[j\pi p(f-f_0)^2] \qquad 3.2.4$$

(Klauder et al., 1960).

The quadratic phase characteristic leads to a linear group delay response given by

$$t_d = -\frac{1}{2\pi}\frac{d}{df}[\pi p(f-f_0)^2] = -p(f-f_0). \qquad 3.2.5$$

The group delay is clearly of the form needed to compress a linear up-chirp containing only positive frequencies. If $f_0 >> B$ it can be shown (Cook, 1960) that this condition is fulfilled to a high degree of accuracy. If $p=1/k$ the delay over the input chirp bandwidth is T. In this case "ideal" pulse compression (illustrated in figure 3-2) occurs. If only the ideal case is considered at this stage, the compression network admittance function can be written as

$$H(f) = \exp[j\pi(f-f_0)^2/k] \; . \qquad 3.2.6$$

The impulse response of the network is found by taking the inverse Fourier transform of H(f), so

$$h(t) = \int_{-\infty}^{\infty} \exp[j\pi(f-f_0)^2/k]\exp[j2\pi ft]df.$$

Re-arrangement of the integrand exponent produces the standard form

$$h(t) = \int_{-\infty}^{\infty} \exp\left\{j2\pi[ft+(f-f_0)^2/2k]\right\}df$$

which may be solved (Klauder et al., 1960) to give

$$h(t) = \sqrt{jk}\,\exp[j2\pi(f_0 t - \tfrac{1}{2}kt^2)] \quad , \qquad 3.2.7$$

so the impulse response of the compression network is a down-chirp.

The output, g(t), of the compression filter is easily deduced using a standard time domain method. By the convolution theorem,

$$g(t) = \int_{-\infty}^{\infty} f(\tau)h(t-\tau)\,d\tau \quad .$$

Substituting from equations 3.2.1 and 3.2.7 gives

$$g(t) = \sqrt{jk}\int_{-\infty}^{\infty} w(\tau)\exp[j2\pi(f_0\tau+\tfrac{1}{2}k\tau^2)]\exp\left\{j2\pi[f_0(t-\tau)-\tfrac{1}{2}k(t-\tau)^2]\right\}d\tau \quad .$$

Re-arrangement and cancellation in the exponent produces the form

$$g(t) = \sqrt{jk}\,\exp[j2\pi(f_0 t-\tfrac{1}{2}kt^2)]\int_{-\infty}^{\infty} w(\tau)\exp[j2\pi kt\tau]d\tau$$

which can be re-written as

$$g(t) = \sqrt{jk}\,\exp[j2\pi(f_0 t-\tfrac{1}{2}kt^2)]W(-kt) \quad , \qquad 3.2.8$$

where W(f) is the Fourier transform of the input chirp weighting function w(t).

Hence, to compress a chirp signal a compression network having a chirp impulse response is required. If the input chirp slope is k Hertz/second then the compression network chirp slope must be -k if ideal compression is to occur. The collapsed output has the form of a chirp (of opposite slope to the input chirp) multiplied by an envelope function which is the Fourier transform of the input chirp weighting function.

A device having a chirp impulse response is called a dispersive delay line (DDL) and, as well as functioning as a compression filter, can be used to generate a chirp signal. A DDL used in the latter role is known as an expander since an impulse (or, in practice, a pseudo-impulse) is expanded in duration into a chirp.

3.3 UN-WEIGHTED CHIRP COMPRESSION

An important special case arises when the envelope of the input chirp is given by the rectangular gating function defined by equation 3.2.3 and sketched in figure 3-1(b). The Fourier transform of such a function is

$$W(f) = T\frac{\sin(\pi fT)}{\pi fT} \quad . \qquad 3.3.1$$

Substituting in equation 3.2.8 and recalling that k=B/T gives

$$g(t) = \sqrt{jTB}\,\exp[j2\pi(f_o t - \tfrac{1}{2}kt^2)]\,\frac{\sin(\pi Bt)}{\pi Bt} \quad . \qquad 3.3.2$$

The envelope of the compressed output has a sin x/x (or sinc x) form characterised by a -4 dB width of 1/B, nulls at ± n/B (n non-zero and integral), and a largest minor response of about -13dB. The output is sketched in figure 3-3. The Fourier transform relationship between the compressor input and output signal envelopes is identical to that relating the aperture illumination function and the far-field response of an antenna. As in the antenna case, the minor responses are known as sidelobes.

The compression gain of the system is $\sqrt{TB}$, so the peak amplitude of the compressed pulse is $\sqrt{TB}$ times the amplitude of the input chirp. This is expected from the energy conservation requirement, since decreasing the time occupied by a signal must increase its amplitude. Section 6.3 illustrates the practical use of this concept.

Although the instantaneous frequency characteristic sketched in figure 3-1(a) is useful in describing the form of chirp signals, Cook (1960) points out that it will not always give an accurate picture of the chirp spectrum. An accurate spectral representation is obtained by taking the Fourier transform of equation 3.2.1 and, for the un-weighted or uniform chirp, the spectrum is given by

$$F(f) = \int_{-T/2}^{T/2} \exp[j2\pi(f_o t + \tfrac{1}{2}kt^2)]\exp(-j2\pi ft)dt \quad .$$

Klauder et al. (1960) show that the normalized amplitude spectrum resulting from the above transformation can be written as

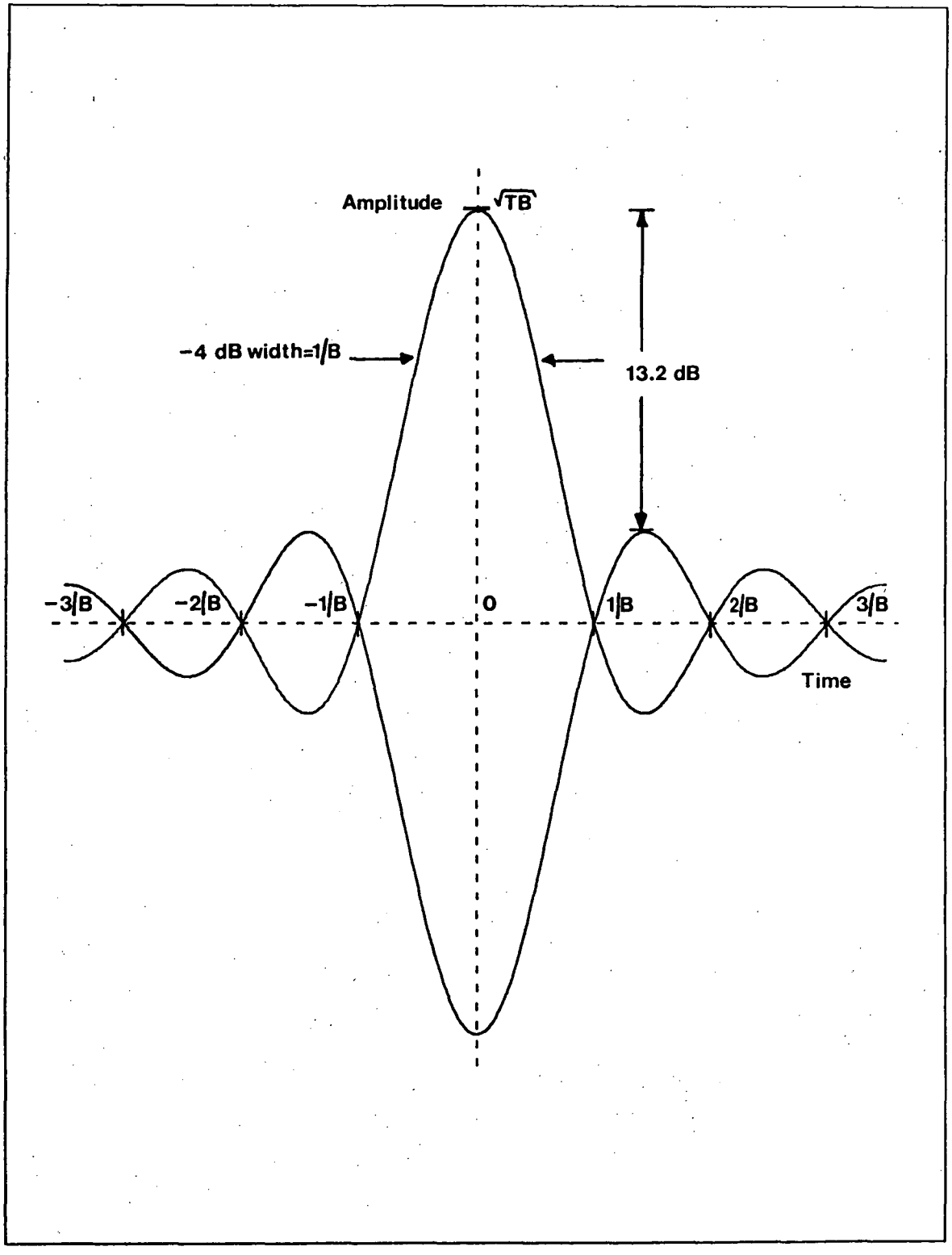


Fig.3-3: The sinc(x) output resulting from the compression of a unit amplitude un-weighted chirp in an ideal group delay equalizer (the RF carrier component is not shown). The -4 dB width is equal to the inverse bandwidth of the compression operation and the compression gain is $\sqrt{TB}$.

$$|F(f)| = \frac{1}{\sqrt{2}}\left\{[C(u_2)-C(u_1)]^2+[S(u_2)-S(u_1)]^2\right\}^{\frac{1}{2}} \quad , \qquad 3.3.3$$

where

$$u_1 = -2(f-f_o)\sqrt{\frac{1}{2k}} - \sqrt{\frac{TB}{2}}$$

$$u_2 = -2(f-f_o)\sqrt{\frac{1}{2k}} + \sqrt{\frac{TB}{2}}$$

and C and S are the real and imaginary parts of the complex Fresnel integral,

$$C(u)+jS(u) = \int_0^u \exp[j\pi x^2/2]dx \quad .$$

The amplitude given by equation 3.3.3 is plotted in figure 3-4 as a function of normalized frequency for various values of time-bandwidth product. Also plotted is the angle

$$\phi_d = -\arctan\left\{\frac{S(u_2)-S(u_1)}{C(u_2)-C(u_1)}\right\} \qquad 3.3.4$$

(adapted from Cook, 1960), representing the deviation from a quadratic phase law. For high values of TB the spectrum approaches the rectangular form expected on the basis of a simple instantaneous frequency model. However, even for TB=10 where the spectrum is badly distorted, almost 95% of the chirp energy falls within the nominal bandwidth.

The chirp "Fresnel ripple" can be understood qualitatively by recalling the duality between time-domain multiplication and frequency-domain convolution. Multiplication of a chirp by a gating function results in a spectrum which is the convolution of the chirp and gating function spectra. The spectrum of a rectangular function has a sinc(x) form and the convolution of this function with the rectangular spectrum of an infinite duration chirp produces a product spectrum exhibiting Fresnel ripple. As the gating function narrows, its spectrum widens and the sinc(x) function tends to dominate in the convolution, leading to more severe distortion of the rectangular spectrum.

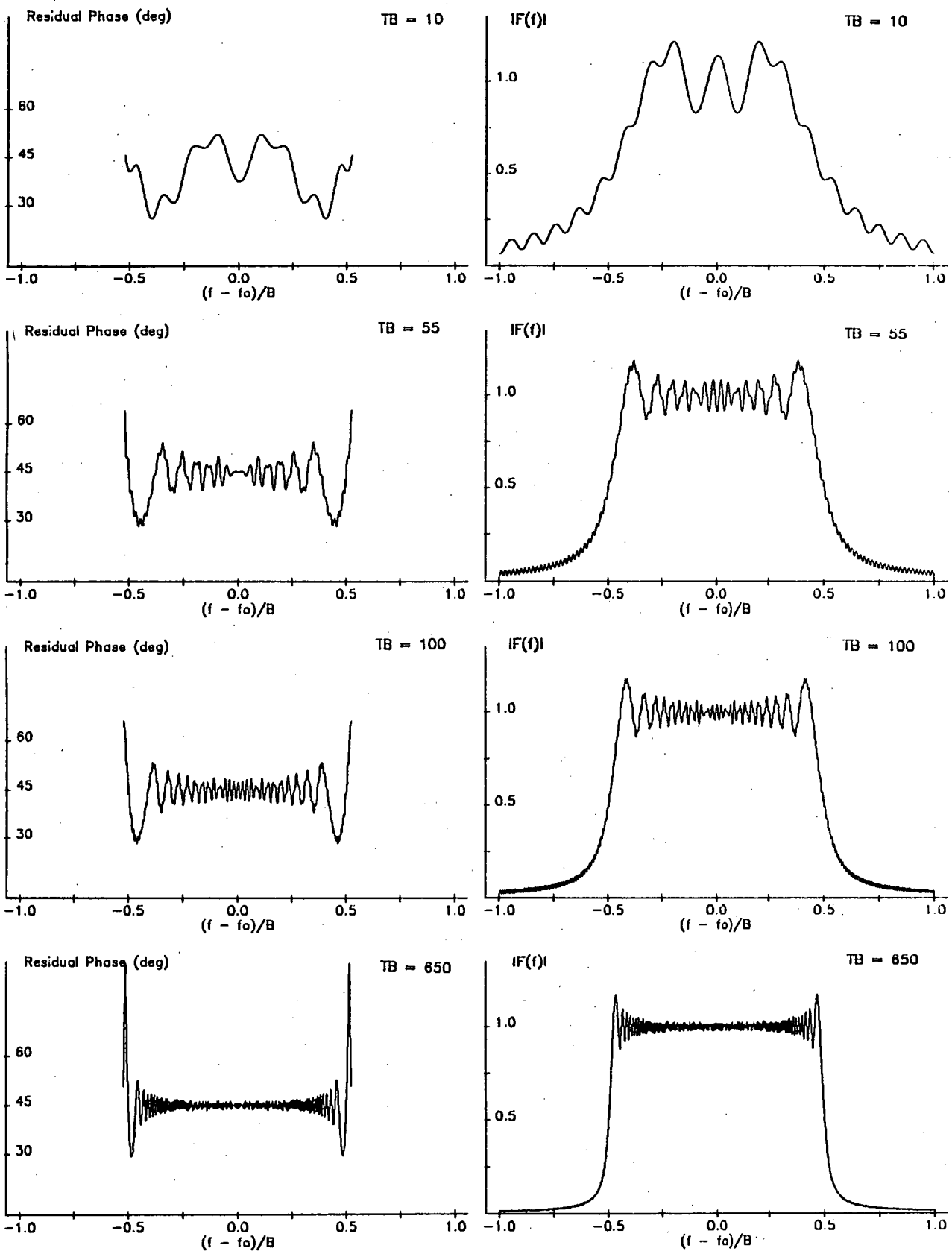


Fig.3-4: Normalized amplitude and residual phase spectra for chirp signals having various values of time-bandwidth product. As TB increases the spectrum shape approaches that predicted using the instantaneous frequency model of figure 3-1(a).

3.4 MATCHED FILTER PULSE COMPRESSION

The pulse compression process can be viewed as an example of matched filtering since the response of the compression filter is matched to the input chirp. In general, the frequency response of a filter matched to an input, f(t), can be shown (see, for example, Schwarz, 1980) to be

$$H_m(f) = F^*(f)\exp(-j2\pi fT_d) \quad , \qquad 3.4.1$$

where $F^*(f)$ is the complex conjugate of the Fourier transform of f(t) and the delay constant, T_d, is required if the filter is to be realizable. The filter response is derived by optimizing the output signal-to-noise ratio at $t=T_d$ in the presence of additive Gaussian noise. As in section 3.2, it is convenient to set $T_d=0$ for analysis purposes.

In the time domain the equivalent matched filter specification is

$$h_m(t) = f^*(-t) \qquad 3.4.2$$

which, for a real input signal, means that the impulse response of a matched filter is the time inverse of the signal being filtered. The time domain convolution integral describing the matched filter output has the same form as the autocorrelation integral for the input signal, f(t). Pulse compression can therefore be considered to be a correlation of the input signal with the chirp impulse response of the compressor.

Consider an un-weighted chirp defined by equations 3.2.1 and 3.2.3. From equation 3.4.2 the matched filter impulse response is given by

$$h_m(t) = \prod\left(\frac{t-\frac{1}{2}T}{T}\right)\exp[j2\pi(f_0t-\tfrac{1}{2}kt^2)] \quad , \qquad 3.4.3$$

which is of the same form as the impulse response of the delay equalizer compression filter (equation 3.2.7) except for the truncation or gating function. The matched filter output can be deduced by convolving the the input signal with the filter impulse response. Thus,

$$g_m(t) = \int_{-\infty}^{\infty} \Pi\left(\frac{\tau-\frac{1}{2}T}{T}\right)\Pi\left(\frac{t-\tau-\frac{1}{2}T}{T}\right) \exp[j2\pi(f_o\tau+\tfrac{1}{2}k\tau^2)]$$

$$\exp\left\{j2\pi[f_o(t-\tau)-\tfrac{1}{2}k(t-\tau)^2]\right\}d\tau$$

which simplifies to

$$g_m(t) = \exp[j2\pi(f_o t-\tfrac{1}{2}kt^2)]\int_{t-\frac{1}{2}T}^{\frac{1}{2}T} \exp[j2\pi kt\tau]d\tau \; ; \; 0\le t\le T.$$

Klauder et al. (1960) show that when the integral is evaluated the output is given by:

$$g_m(t) = T\exp(j2\pi f_o t)\frac{\sin[\pi(B|t|-kt^2)]}{\pi B|t|} \; ; \quad |t|\le T$$

and

$$g_m(t) = 0 \qquad\qquad ; \quad |t|>T.$$

Substituting y = Bt and normalizing to permit direct comparison with equation 3.3.2 gives

$$g_m'(t) = \sqrt{TB}\exp(j2\pi f_o t)\frac{\sin[\pi\left(|y| - \frac{y^2}{TB}\right)]}{\pi|y|} \; ; \; |t|\le T. \qquad 3.4.4$$

The normalized envelope of this function is plotted for several values of time-bandwidth product in figure 3-5. For large values of TB the output of a matched filter resembles closely the sinc(x) output of an infinite bandwidth group delay equalizer although the carrier residual FM implied in equation 3.3.2 is not present. Note that the matched filter output lies within the triangle resulting from the autocorrelation of the rectangular input envelope.

For chirps of low time-bandwidth product the spectrum exhibits significant Fresnel ripple (figure 3-4) and the matched filter requirement (equation 3.4.1) dictates the need for a filter with equally complicated amplitude and phase characteristics. The similarity of the matched filter and the delay equalizer responses means that, in practice, a compressor with idealized characteristics (uniform amplitude and quadratic phase) but having a truncated response in time and frequency can be used.

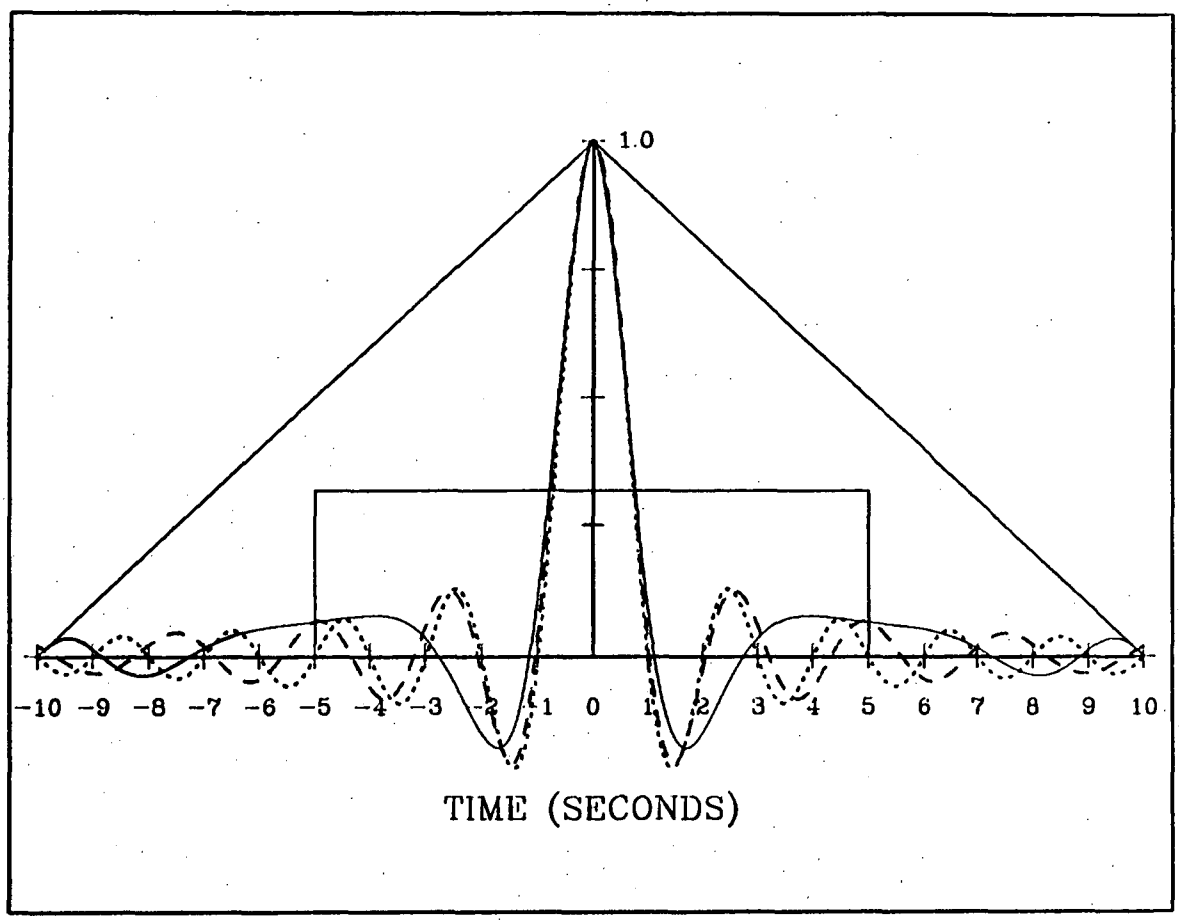


Fig.3-5: Normalized envelope of a matched filter output for three values of time-bandwidth product. All solid lines relate to TB = 10, the dashed line is for TB = 55 and the dotted line represents a sinc(x) (infinite TB) output. As well as the amplitude normalization, the bandwidth is set to unity in each case. The rectangle is the envelope of an un-weighted chirp (TB = 10) and the triangle is the envelope autocorrelation function. Notice that the TB = 10 output is bounded by the autocorrelation triangle.

In recent years the adoption of SAW technology has made it possible to realise very complicated responses and close approximations to true matched filters are now often used as compressors.

3.5 CHIRP WEIGHTING AND DDL RESPONSE ACCURACY

The first sidelobe level of a sinc(x) pulse is about -13 dB (figure 3-3). In most applications this level is unacceptable and chirp weighting schemes analogous to antenna aperture tapering are used to improve the sidelobe suppression. As well as sidelobe reduction, weighted compression reduces the compression loop sensitivity to chirp slope mismatch between the expander and compressor dispersive filters. The penalties for weighting are an increase in compressed pulse main lobe width and, since the compressor is no longer matched exactly to its input, a small reduction in compression gain. For the spectrum analyser application (figure 2-9), weighting degrades the frequency resolution but improves the dynamic range since weak signals are not masked by sidelobes from very strong signals. In radar applications weighting improves the visibility of weak target returns in the presence of adjacent strong returns at the expense of degrading the range resolution.

The effects on the compressed pulse of weighting or of phase and amplitude response errors in a compression system can be estimated using paired echo theory (Wheeler, 1939). The theory considers the phase and amplitude characteristics of a linear network to be composed of an ideal response (flat amplitude, linear phase) plus error components. The error components modulate the main response leading to the production of the time domain equivalent of sidebands. These secondary time responses are known as echoes and may lead or lag the main output. If, for example, the amplitude-frequency distortion of the network can be described by the simplest modification of the ideal response, a single cosinusoidal (even) ripple characteristic, the effect is to produce a leading and lagging echo pair having the same polarity as the main output. Similarly, for a small sinusoidal (odd) variation of the nominally odd linear phase function the echo pair is anti-symmetrical.

Cook and Bernfeld (1967, ch. 11) describe in detail the application of paired echo analysis to pulse compression loops having

phase and amplitude distortion. A convenient method of analysis is to consider the system to be ideal but for the presence of a fictitious filter which modifies the chirp spectrum prior to compression. Distortion in the compression loop, apart from that resulting from chirp slope mismatch, is ascribed to deviation from a flat amplitude or linear phase response in the fictitious filter.

Assuming sinusoidal response modulation, an amplitude error of 0.085 dB or a phase error of 0.57^{o} is sufficient to generate a spurious response (or coherent noise) level of -40 dB relative to the compressed pulse peak. Deviations of 0.6 dB or 4^{o} set the coherent noise level at -30 dB. Compression systems using SAW technology can yield spurious response levels of less than -45 dB although -35 dB is more common if pre-distortion methods are not employed in the dispersive filter design.

It is important to realise that the effects of response distortion are not, in general, removable by the frequency or time domain weighting methods to be described. Indeed, large amounts of distortion can negate the effects of weighting. Gerard et al (1973) discuss the effect of random phase errors in dispersive devices and confirm the predictions of paired echo theory that small phase errors can produce time echoes large enough to reduce the effectiveness of sidelobe reduction schemes. In one weighted SAW dispersive delay line (TB=1000) it was found that a non-periodic phase error of 20 degrees produced a sidelobe level of greater than -27 dB instead of the -42 dB level predicted by weighting theory.

The most obvious method of chirp weighting is to make the gating function w(t) (equation 3.2.1) non-uniform. From equation 3.2.8, the compressed pulse envelope then departs from a sinc(x) form. An expression similar to equation 3.2.8 can be derived for a matched filter compressor. If w(t) is Gaussian for example, a first sidelobe level of -35 dB is obtained and the pulse broadening factor (defined at the -4 dB points) relative to a sinc(x) response is 1.9. Antenna design handbooks give similar information for other weighting functions (see also figure 3-8).

In practice, it is often undesirable to compress a weighted incoming chirp. In a chirp spectrum analyser generation of properly weighted signal chirps requires bi-linear operation of the signal mixer. For wideband analysers this requirement is impossible to meet. Radar

sensitivity requirements make it essential to transmit as much energy as possible, so for a given pulse length a uniform chirp is also desirable in this application.

The usual way of producing the effect of a weighted chirp is to taper the amplitude-frequency response of the compressor DDL. The duality of time and frequency in chirp signals (exact for infinite TB) means that this process is identical in concept to one of modifying the amplitude-time characteristic of the input chirp. Taking an approach similar to that outlined in the discussion of system response errors, it is convenient to consider a weighting filter as a frequency response shaping device (with a linear phase characteristic) inserted external to the compressor. The compression filter may be either a group delay equalizer or a matched filter.

The sidelobe reduction effect of a weighting network is readily explained with the aid of paired echo theory. If, for example, the weighting filter has a raised cosine response (figure 3-6) the combination of the main output with leading and lagging echoes (displaced by 1/B - the inverse of the cosine spectral period) produces a pulse with a broader central lobe but reduced sidelobes (figure 3-7). The raised cosine taper is a particular example of a class of weightings having the form

$$W(f) = P + (1-P)\ \cos^n \left\{\frac{\pi(f-f_o)}{B}\right\} \qquad 3.5.1$$

and known as cosine power weightings. The choice of n=2 and P=0.08 produces the well-known Hamming weighting function, leading to far sidelobes (those outside the region ± T/2) which fall off at the rate of 6 dB/octave in time (McCue, 1980). For n=2, the Hamming pedestal height of P=0.08 produces optimum near sidelobe suppression. Cook and Bernfeld (1967, ch.7) demonstrate that sidelobe suppression is much more sensitive to changes in pedestal height than to variations in the shape of the harmonic weighting function.

The weighting function giving the lowest sidelobe level consistent with the smallest increase in pulse width is the Dolph-Tchebycheff function. The response is not realizable physically but serves as a standard for comparison. The Taylor weighting function

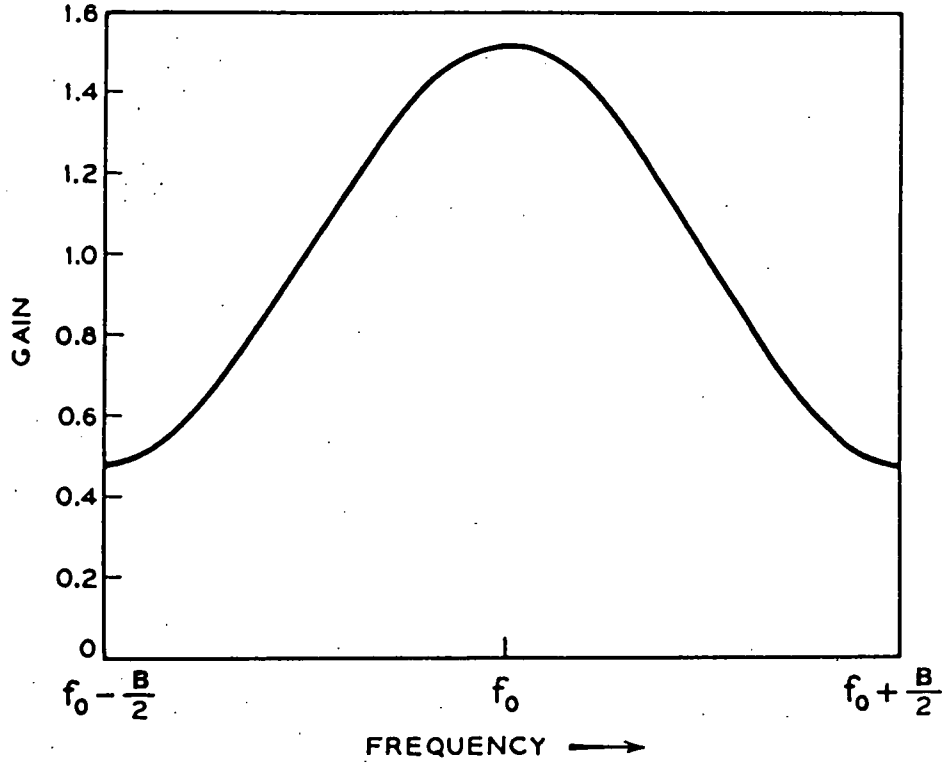


Fig.3-6: Frequency response of a raised cosine chirp weighting filter. (After Klauder et al., 1960).

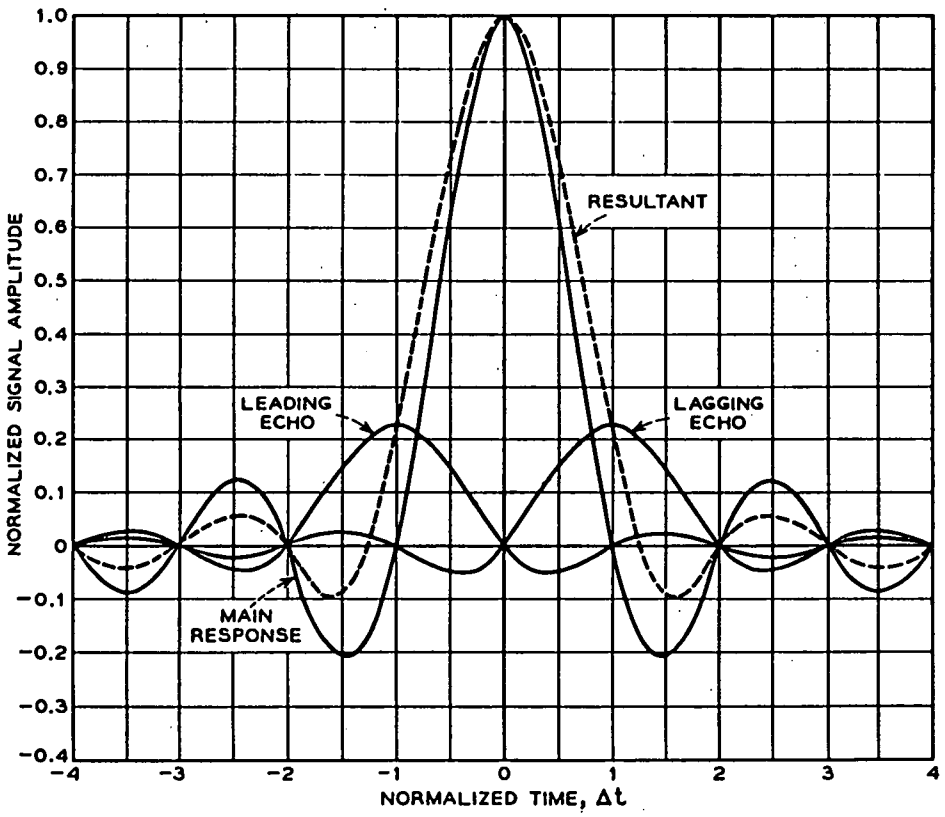


Fig.3-7: The sidelobe reduction effect of a chirp weighting network. Combining the main response with the leading and lagging echoes produces a pulse having a broader main lobe but reduced sidelobes. (After Klauder et al., 1960).

(Taylor, 1955) is often used as an approximation to the Dolph-Tchebycheff response. A parameter of the Taylor weighting, usually denoted by $\bar{n}$, effectively specifies the closeness of the approximation. Figure 3-8 shows the trade-off between main lobe and sidelobe level for several weighting functions.

As mentioned earlier, time domain weighting of the chirp to be compressed is usually undesirable in practice. However, reference to figure 3-7 suggests an alternative post-compression time domain weighting arrangement. The sidelobe reduction effect can be achieved by adding appropriately scaled and delayed versions of the main pulse. A possible realization of such a system is the transversal filter illustrated in figure 3-9. At least one practical implementation of this method has been reported (Campbell et al., 1982) and the scheme is expected to be employed in the spectrum analysers described in appendix five.

This brief discussion of weighting has, like most literature on compression techniques, assumed that the compression filter output is a sinc(x) function. Matched filter compression results in a pulse which departs significantly from a sinc(x) form at low values of time-bandwidth product (figure 3-5). In such cases the peaks and troughs in the wings of the pulse are not evenly spaced. McCue (1980) has pointed out that weighting methods giving secondary time responses having fixed delays relative to the main output do not result in sidelobe cancellation in the region $\pm$ T/2. These Fresnel or gating sidelobes, having a level of about 20 log(TB)+3 dB below the compressed pulse peak (Andersen Laboratories, 1981) may be troublesome in compression systems with a low time-bandwidth product.

Judd (1973) has described a "reciprocal ripple compensation" design technique for the realization of low time-bandwidth product compressors exhibiting superior sidelobe characteristics. In a large TB compression loop the compressed pulse spectrum is very similar to the frequency response of the weighting network. If a low TB compressor incorporating a weighting network can be designed such that the compressed pulse spectrum is similar to that obtained with a high TB device, the sidelobe performance should be similar to that obtained from the large TB process (Butler, 1979).

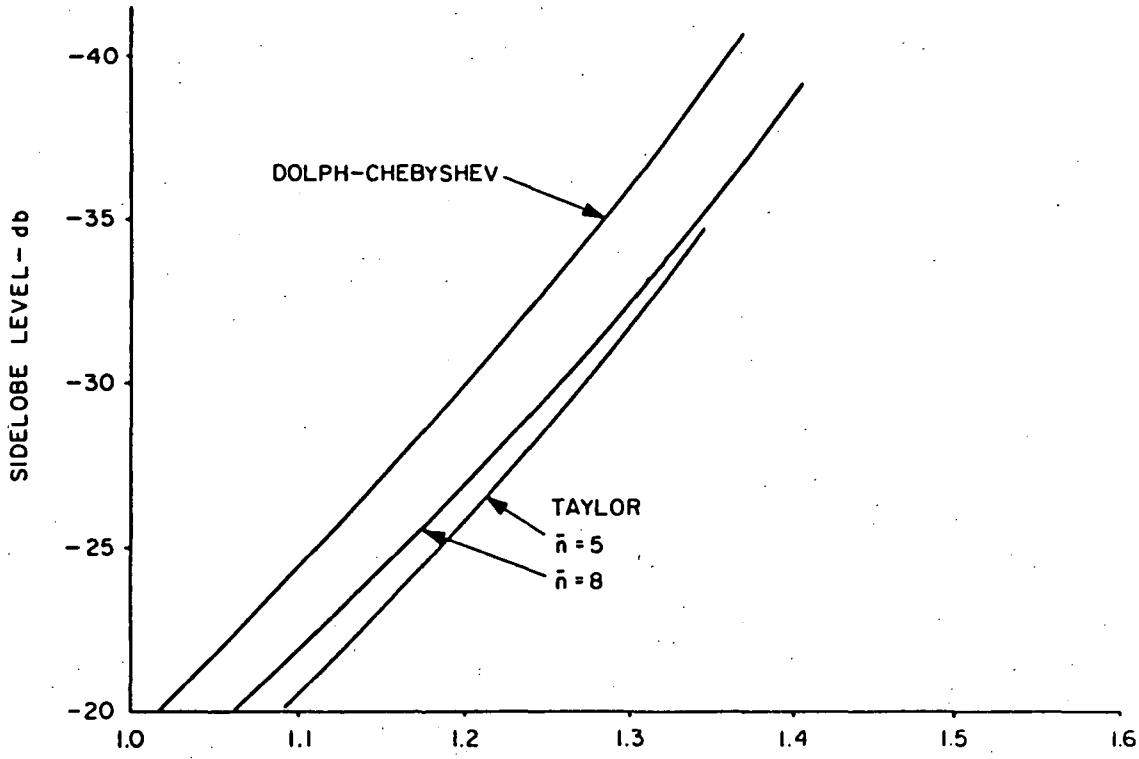


Fig.3-8: The increase in main lobe width necessary to achieve a specified sidelobe level with three weighting functions. (After Cook and Bernfeld, 1967, ch.7).

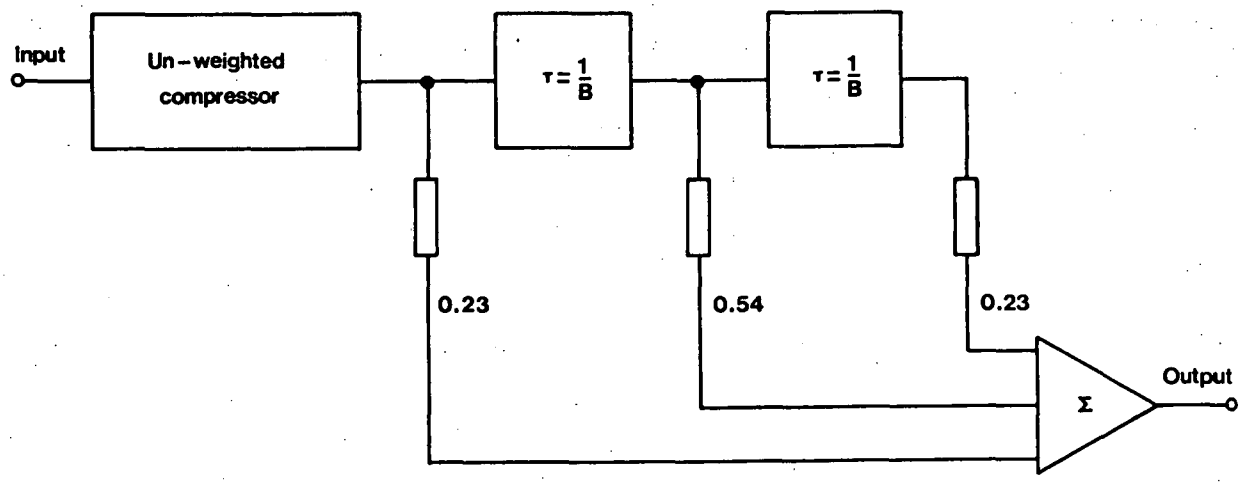


Fig.3-9: Post-compression weighting network based on a transversal filter. The coefficients shown result in an approximation to the Hamming weighting function.

If the required spectrum is G(f) and the input chirp spectrum if F(f), the required compressor response is

$$H(f) = G(f)/F(f) \quad .$$

In this equation F(f) appears in the denominator so Fresnel ripple in the input chirp appears as "reciprocal ripple" in the compressor response. SAW technology allows filters with the required response complexity to be realized.

For the reciprocal ripple technique to be effective the centre frequency of the chirp to be compressed must be constant and its Fresnel ripple characteristics must match the compressor response. Depending on the form of chirp spectrum analyser chosen (chapter four), one or both of these conditions will not apply and so the technique has little application in spectrum analyser dispersive filters. In radar systems reciprocal ripple compressors exhibit high sidelobe sensitivity to Doppler frequency shift on chirps reflected from moving targets, again limiting the usefulness of the method.

Although reduction of time sidelobes is the main reason for utilising a weighting network, its inclusion reduces distortion of the compressed pulse when the slope of the input chirp is not matched exactly by that of the compression network. This type of imperfection is referred to as quadratic phase distortion and is to be distinguished from the random or partially random phase errors discussed earlier.

It can be shown (Klauder et al., 1960) that even when the quadratic phase variation between the signal and the compressor gives rise to phase departures of 360^{o} at the band edges, heavily weighted systems typically produce a single pulse exhibiting a loss of 4 dB and a main lobe broadening factor of about 2 relative to the ideal output. The sidelobe level remains at around -35 dB. In contrast, an un-weighted compressor produces a dual-peaked pulse showing a 7 dB loss, a ten-fold increase in width and sidelobe levels of only -6 dB. Cook and Bernfeld (1967, ch.6) show that, for a given chirp slope mismatch factor, pulse widening depends on time-bandwidth product with large TB systems producing greater widening factors.

CHAPTER 3 - REFERENCES

Andersen Laboratories: Dispersive delay lines: matched filters for signal processing. Technical note, 1981.

Butler, M.B.N.: SAW devices for signal processing. IEE Conf. Pub.180, 1979, pp.21-33.

Campbell, P.A., Grant, P.M. and Arthur, J.W.: Extension of time-bandwidth product in SAW spectrum analysers. Electron. Lett., Vol.17, No.16, Aug. 1981, pp.576-578.

Cook, C.E.: Pulse compression: Key to more efficient radar transmission. Proc. IRE, Vol.48, March 1960, pp.310-316.

Cook, C.E. and Bernfeld, M.: "Radar signals: an introduction to theory and application". Academic Press, 1967.

Gerard, H.M., Smith, W.R., Jones, W.R. and Harrington, J.B.: The design and applications of highly dispersive acoustic surface wave filters. IEEE Trans. Son. Ultrason., Vol.SU-20, No.2, April 1973, pp.94-104.

Judd, G.W.: Technique for realizing low time sidelobe levels in small compression ratio chirp waveforms. Proc. 1973 IEEE Ultrasonics Symp., pp.478-481.

Klauder, J.R., Price, A.C., Darlington, S. and Albersheim, W.J.: The theory and design of chirp radars. Bell Syst. Tech. J., Vol.39, No.4,July 1960, pp.745-808.

McCue, J.J.G.: Suppression of range sidelobes in bistatic radars. Proc. IEEE, Vol.68, No.3, March 1980, pp.422-423.

Schwarz, M.: "Information transmission, modulation, and noise". Ch.5, McGraw-Hill Kogakusha, 1980.

Taylor, T.T.: Design of line source antennas for narrow beamwidths and low sidelobes. Trans. IRE, Vol. AP-3, No.1, 1955, pp.16-28.

Wheeler, H.A.: The interpretation of amplitude and phase distortion in terms of paired echoes. Proc. IRE, Vol.27, June 1939, pp.359-383.

CHAPTER 4 SAW SPECTRUM ANALYSERS

4.1 INTRODUCTION

In chapter two a brief description of the compressive or microscan receiver spectrum analyser was given. The style of that description paralleled the historical development of the instrument: an IF filter having a non-constant group delay was substituted for the usual linear phase (flat group delay) filter in a standard scanning spectrum analyser. The substitution enabled the analyser to exhibit significantly better frequency resolution than a standard analyser having the same LO scan rate.

In this chapter the continuous chirp Fourier transform will be reviewed and the microscan receiver will be shown to be a simplified implementation of one of two possible chirp transform algorithms. The result of both chirp transformations is a duality between frequency at the Fourier transformer input and time at its output. The pulse compression process occurring within the transformer is the same, to the first order, as that in the basic compression loop discussed in chapter three.

Chapter four is divided into two sections. Part A is a review of SAW spectrum analyser theory and basic configurations. The review is one of the most complete written and includes references to newly completed or recently de-classified reports dealing with SAW analyser performance (Chomiki, 1984; Davie and Arthur, 1979). As well as the review material, section 4.4 explains the input frequency dependent sidelobe structure in low time-bandwidth product analysers. Part B of the chapter deals with topics related to analog and digital post-transformation signal processing. A substantial fraction of the material presented in the second part of the chapter is original although references to previous work are made where appropriate.

Before beginning a description of analog chirp transformation it is appropriate to review the current status of SAW processors. The main advantage of the SAW implementation of the chirp transform is the speed with which wideband spectral analysis can be performed. Present technology allows the manufacture of RAC DDL's with TB~10 000 and current upper limits on bandwidth and time are about 500 MHz and 100 μs respectively. Thus, as an example, a bandwidth of 500 MHz

can be analysed in 20 μs with 50 kHz frequency resolution. The speed attribute has led to the widespread adoption of SAW technology by military users in applications where high time resolution spectrum analysis means a high probability of intercepting sporadic signals (Nichols, 1982).

The nature of the spectrum readout has largely prevented the use of SAW spectrum analysers in applications where extensive post-transformation processing is needed. The fast serial analog output means that any post-transformation processor (digital or analog) must be capable of processing (or at least acquiring) the emerging spectral data at a very high speed. More quantitatively, the minimum processor bandwidth or clock frequency must be of the order of the bandwidth being analysed by the SAW instrument. With current technology this requirement is difficult to meet for instruments designed to analyse more than a few megahertz.

Prior to the present study no attempt has been made to use SAW instrumentation in radioastronomy although subsequent to the submission of a description of the prototype SAW polarimeter for publication (Hall et al., 1982), another author has pointed to the potential of SAW technology in this field (Morgan, 1982). Alldritt et al. (1978) and Oliver (1978, 1980) have used a narrowband SAW spectrum analyser as part of a military photon counting system but other attempts at signal processing using SAW analysers have been relatively simple ones undertaken as part of military signal intercept schemes.

PART A: ARCHITECTURE AND PERFORMANCE OF CHIRP TRANSFORMERS

4.2 BASIC CHIRP TRANSFORM CONFIGURATIONS

To allow compactness in writing Fourier integrals the angular frequency variable, $\omega = 2\pi f$, will be used in this section. The chirp slope in the angular frequency - time plane is given the symbol μ and is defined such that the instantaneous angular frequency is either μt (for baseband signals) or $\omega_0 + \mu t$ (at RF). As in chapter three the time variable, t, is defined over the range $-T/2 \leq t \leq T/2$, where T is the chirp duration.

The forward Fourier transformation has the form

$$F(\omega) = \int_{-\infty}^{\infty} f(t)\exp(-j\omega t)dt \ . \qquad 4.2.1$$

The full mathematical theory of chirp signal processing has been expounded by Jack (1978) and Jack et al. (1980). One form of the continuous chirp transform algorithm can be illustrated by considering the identity

$$g(t) = \exp(-j\tfrac{1}{2}\mu t^2) \int_{-\infty}^{\infty} f(\tau)\exp(-j\tfrac{1}{2}\mu\tau^2)\exp[j\tfrac{1}{2}\mu(t-\tau)^2]d\tau$$

$$\equiv \int_{-\infty}^{\infty} f(\tau)\exp(-j\mu t\tau)d\tau \ , \qquad 4.2.2$$

(Roberts et al., 1980).

Equation 4.2.2 is functionally the same as the Fourier transform if the transform variable is identified as $\omega = \mu t$; that is, if there is a linear relationship between frequency and time. Equation 4.2.2 suggests therefore that Fourier transformation of a signal, f(t), can be performed by multiplying the signal by a chirp, $\exp(-j\frac{1}{2}\mu t^2)$, convolving the product with an impulse response $\exp(j\frac{1}{2}\mu t^2)$, then multiplying the output signal of the convolver by a third chirp, $\exp(-j\frac{1}{2}\mu t^2)$. In all three operations the numerical value of the chirp slope is the same. The symmetry of the forward and reverse transforms means that chirp methods can produce the inverse transform if the sign of the constituent chirp slopes is chosen appropriately.

The chirp transform algorithm described is known as the multiply-convolve- multiply (or M-C-M) algorithm and can be implemented using SAW chirp filters and standard RF mixers. A block diagram of the resultant Fourier transformer is shown in figure 4-1 (the significance of the chirp durations is discussed in the next section). Processing takes place at RF rather than baseband, so chirps of the form $\exp\{j(\omega_0 + \frac{1}{2}\mu t^2)\}$ are generated. Equation 4.2.2 therefore defines the processing operation with the factor $\exp(j\omega_0 t)$ suppressed throughout.

RF Fourier processors offer some advantage over baseband analog chirp transformers (such as charge coupled device systems) in that

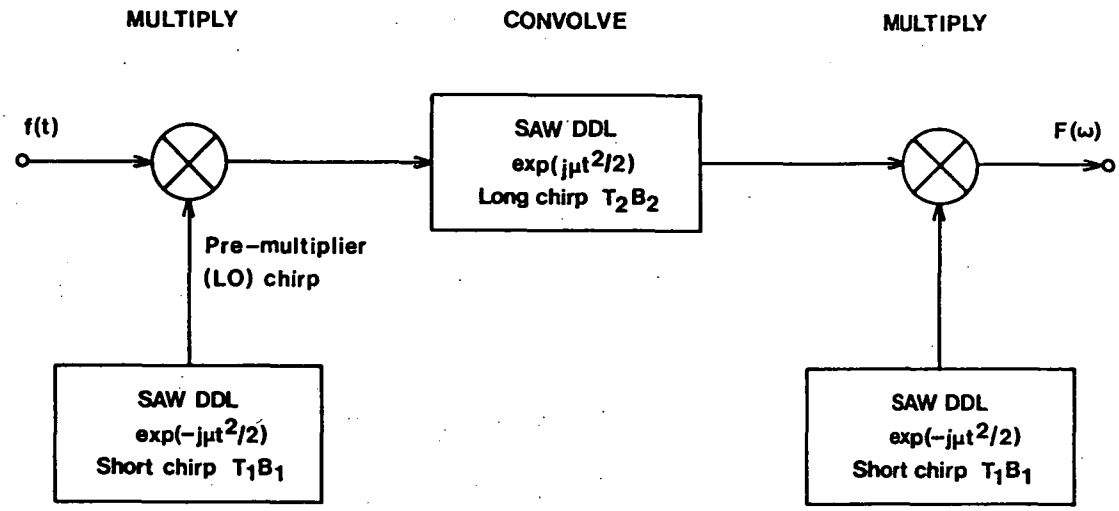


Fig.4-1: M-C-M Fourier transformer based on SAW chirp filters. The multiplication signals are derived by impulsing the appropriate dispersive delay lines.

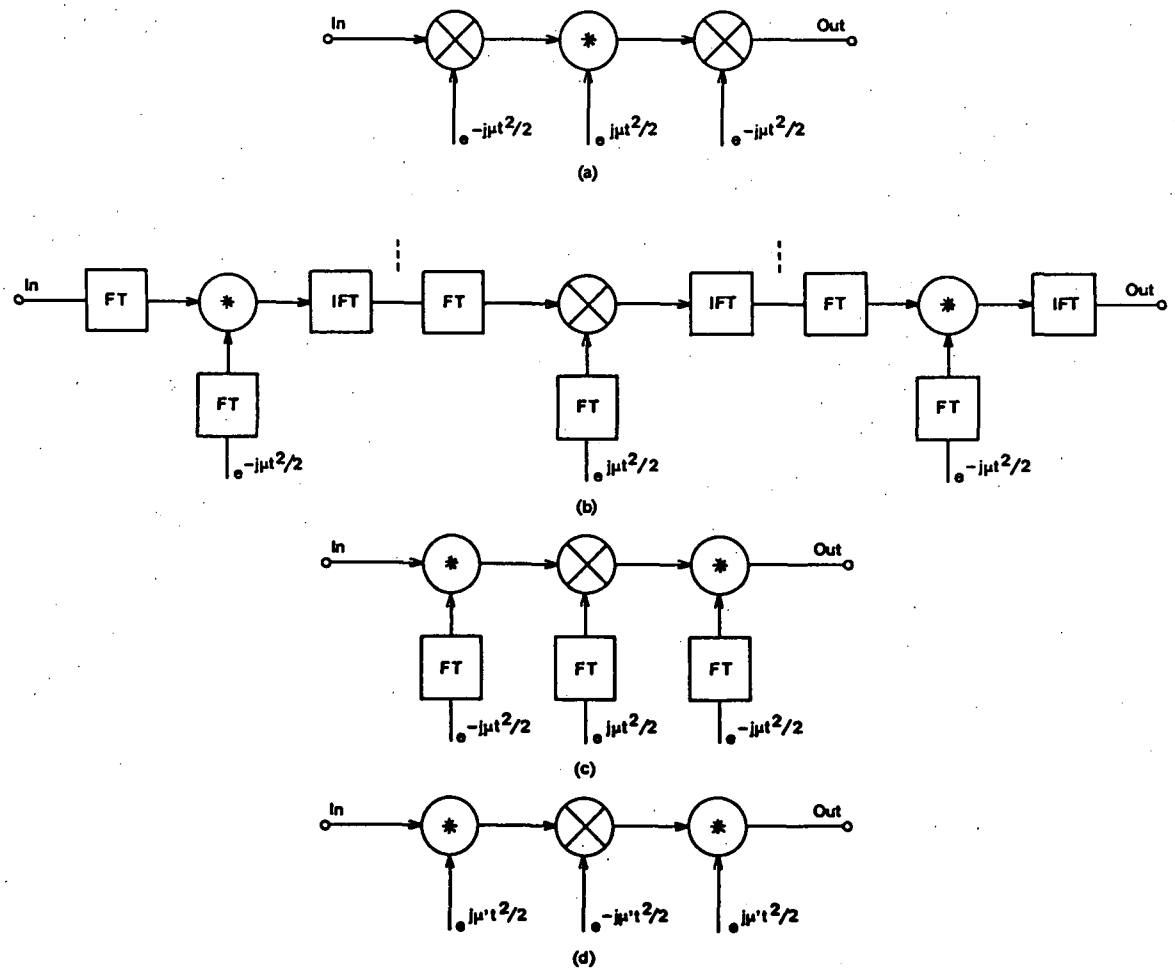


Fig.4-2: Derivation of the C-M-C Fourier transform algorithm from the M-C-M algorithm. The symbol '*' denotes convolution, 'FT' means Fourier transform and 'IFT' means inverse Fourier transform.

harmonic and intermodulation distortion products tend to be displaced widely from the wanted frequency components and are thus able to be filtered effectively. However, in coherent spectrum analysers the RF approach means that system timing is more critical because errors in DDL excitation times (which may be measured in fractions of an RF period) translate directly into phase errors at the analyser output.

The output from the chirp Fourier transformer illustrated in figure 4-1 contains valid phase and amplitude information. If amplitude or power alone is sufficient, the final multiplication (or synchronous detection if the frequency of the post-convolver chirp is chosen appropriately) may be omitted and the convolver output envelope detected using a linear or square-law device. If a complex spectrum is required three DDL's are needed since, although the pre-convolution and post-convolution chirps in the full M-C-M analyser are of the same form, the absolute delay through the convolver DDL means that separately generated chirps are necessary.

The spectral information emerging from the analyser can be displayed on an oscilloscope if the CRO is triggered by successive excitation impulses applied to the LO DDL. The delayed timebase can then be adjusted to compensate for the absolute analyser delay and the CRO time axis read as frequency, the relationship being $\omega = \mu t$.

If the expanded version of equation 4.2.2 is itself Fourier transformed, the "dual" of the M-C-M algorithm can be derived. This can be done diagrammatically by replacing time domain multiplication stages with frequency domain convolution stages and vice-versa. Figure 4-2 illustrates the derivation of the convolve-multiply-convolve (or C-M-C) chirp transform algorithm. The first and last Fourier transforms in (b) are omitted because time and frequency variables can be interchanged. The final simplification, shown in (d), arises because the Fourier transform of an infinite TB chirp with slope μ is a second chirp with slope $-(1/\mu)$. Davie and Arthur (1979) appear to be alone in recognizing the approximation inherent in (d) when chirp signals having finite TB are used. Previous comments relating to the SAW RF implementation of the M-C-M transformer also apply to the C-M-C arrangement.

As well as the two Fourier transform algorithms, several other integral transforms have been implemented using chirp signals. Arsenault and Das (1977) describe a SAW Fresnel transformer and Das and Shklarsky (1979) give details of a real-time Hilbert transform system suitable for producing SSB modulation.

4.3 M-C-M and C-M-C Processors

The two chirp transform algorithms produce, in principle, correct and identical Fourier transformations. In real chirp transform analysers various distortion effects and differences between the two algorithms arise because of the finite extent in time and frequency of the analyser chirp signals.

With finite chirps, the operation of each type of analyser can be illustrated using frequency-time diagrams. For example, figure 4.3(a) shows the operation of the M-C-M algorithm. Although such diagrams are a convenient way of visualizing the processes involved in chirp transformation, they rely on the instantaneous frequency model of chirp signals (section 3.2). Fresnel effects due to the finite time-bandwidth product of real signals are therefore disregarded. In analysers using chirp signals with small values of TB, ignoring these effects leads to an optimistic prediction of compressed pulse sidelobe levels and hence processor dynamic range. However, frequency-time diagrams are helpful in determining fundamental processor parameters such as the relative durations of various chirp signals, the operating duty cycle and the input and output bandwidths.

In the case of an M-C-M Fourier analyser (figure 4-3(a)), the product of the signal and the pre-convolution (or pre-multiplier) chirp must fall within the processing bandwidth of the convolver if valid transformation is to occur. To transform a range of input frequencies the bandwidth of the convolver chirp must therefore exceed that of the pre-multiplier chirp. Since both chirps have the same numerical slope, this requirement means that the pre-multiplier chirp duration must be less than the duration of the convolver impulse response.

Partial transformation occurs on input signals located on either side of the nominal processing bandwidth (figure 4.3(b)). These signals

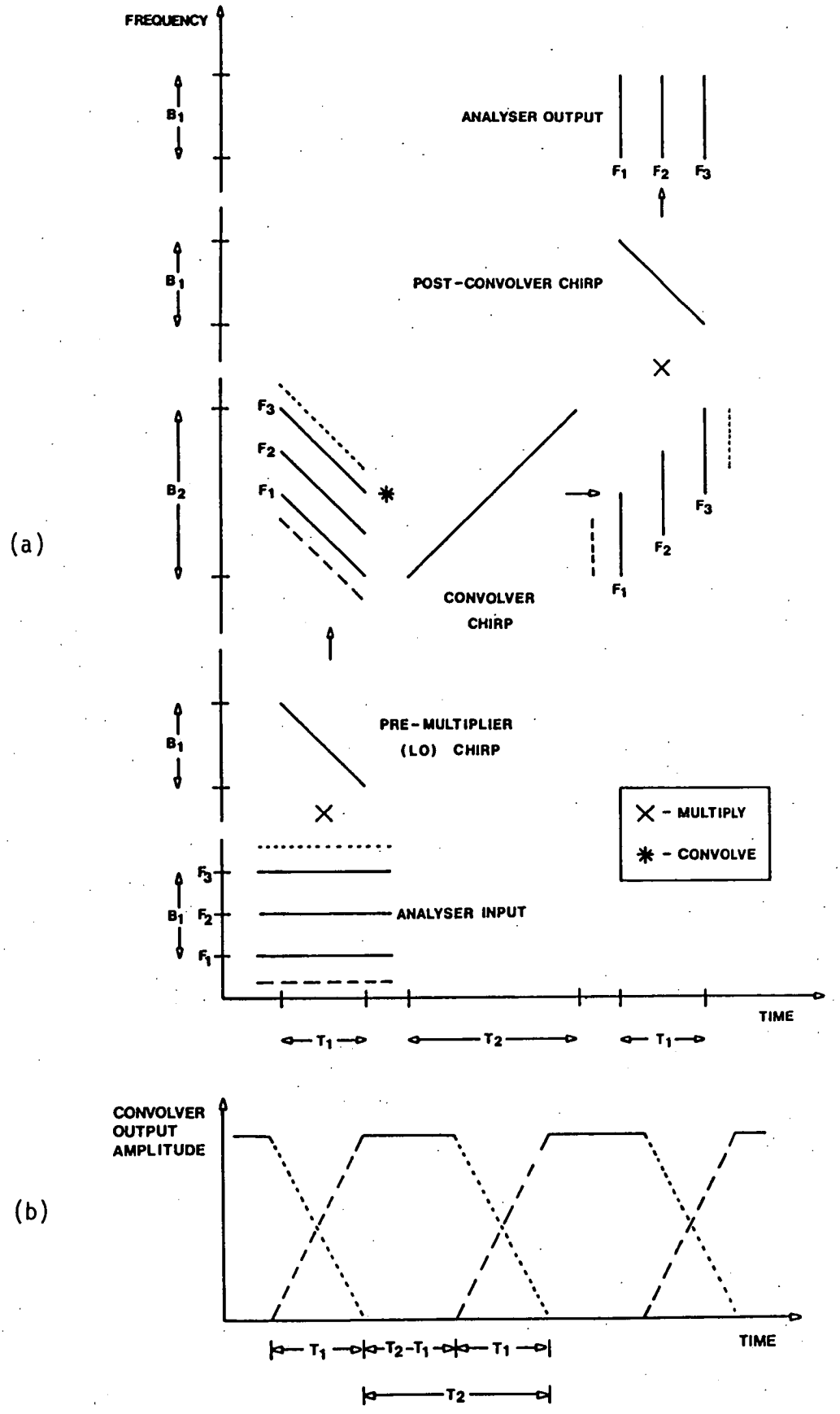


Fig.4-3: Operation of the M-C-M transform algorithm. In (a) a frequency-time diagram for the process is shown. The broken lines represent two out-of-band signals. The idealized frequency response is shown in (b) with the solid lines indicating the full-compression bandwidth and the broken lines the response at the convolver output to out-of-band signals. After the convolver output is multiplied by the final chirp signal, only the full-compression band appears at the analyser output. Successive operating cycles of the analyser are illustrated.

produce chirps which are incompletely compressed, leading to convolver outputs having diminished amplitude and inferior sidelobe characteristics compared with compressed in-band signals.

The analysis bandwidth of the spectrum analyser is the difference in bandwidth of the pre-multiplier and convolver chirps (figure 4-3; Moule, 1982). The duration of the input sample is the pre-multiplier chirp duration. The time-bandwidth product of the input sample is therefore

$$(TB)_i = T_1(B_2 - B_1)$$

where the subscripts 1 and 2 refer to the short and long chirps respectively. If $T_2/T_1 = B_2/B_1 = R$ then

$$(TB)_i = \frac{T_2}{R}\left(B_2 - \frac{B_2}{R}\right)$$

or

$$(TB)_i = B_2T_2\left(\frac{1}{R} - \frac{1}{R^2}\right) \quad . \qquad 4.3.1$$

This expression is maximized when $R = 2$ or $T_1B_1 = T_2B_2/4$, so the most effective processing in an M-C-M analyser occurs when the convolver impulse response chirp is twice the duration of the pre-multiplier chirp. The post-convolver multiplier chirp (if used) need only be as long as the pre-multiplier chirp since it serves only to frequency translate the analyser output over the time window in which full-compression occurs.

For the C-M-C analyser similar reasoning leads to the conclusion that the optimum time-bandwidth product ratio is again four but in this configuration the convolution chirps are the shorter (figure 4-4). A frequency-time diagram for the C-M-C analyser is shown in figure 4-5. A general observation (Moule, 1982) is that in either transform algorithm the central operator (either multiplication or convolution) must be associated with the larger time-bandwidth product.

Otto (1976) has shown that SAW processors, like other Fourier transformers, produce an output corresponding to the actual input spectrum convolved with an instrument resolution function. Over the

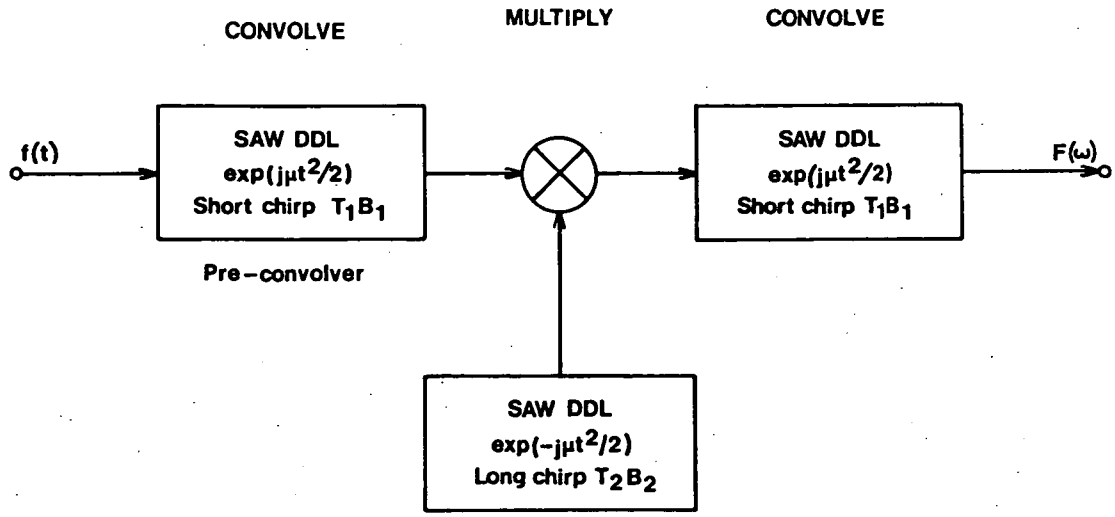


Fig.4-4: C-M-C Fourier transformer using SAW chirp filters. The multiplication signal is derived by impulsing the central dispersive delay line.

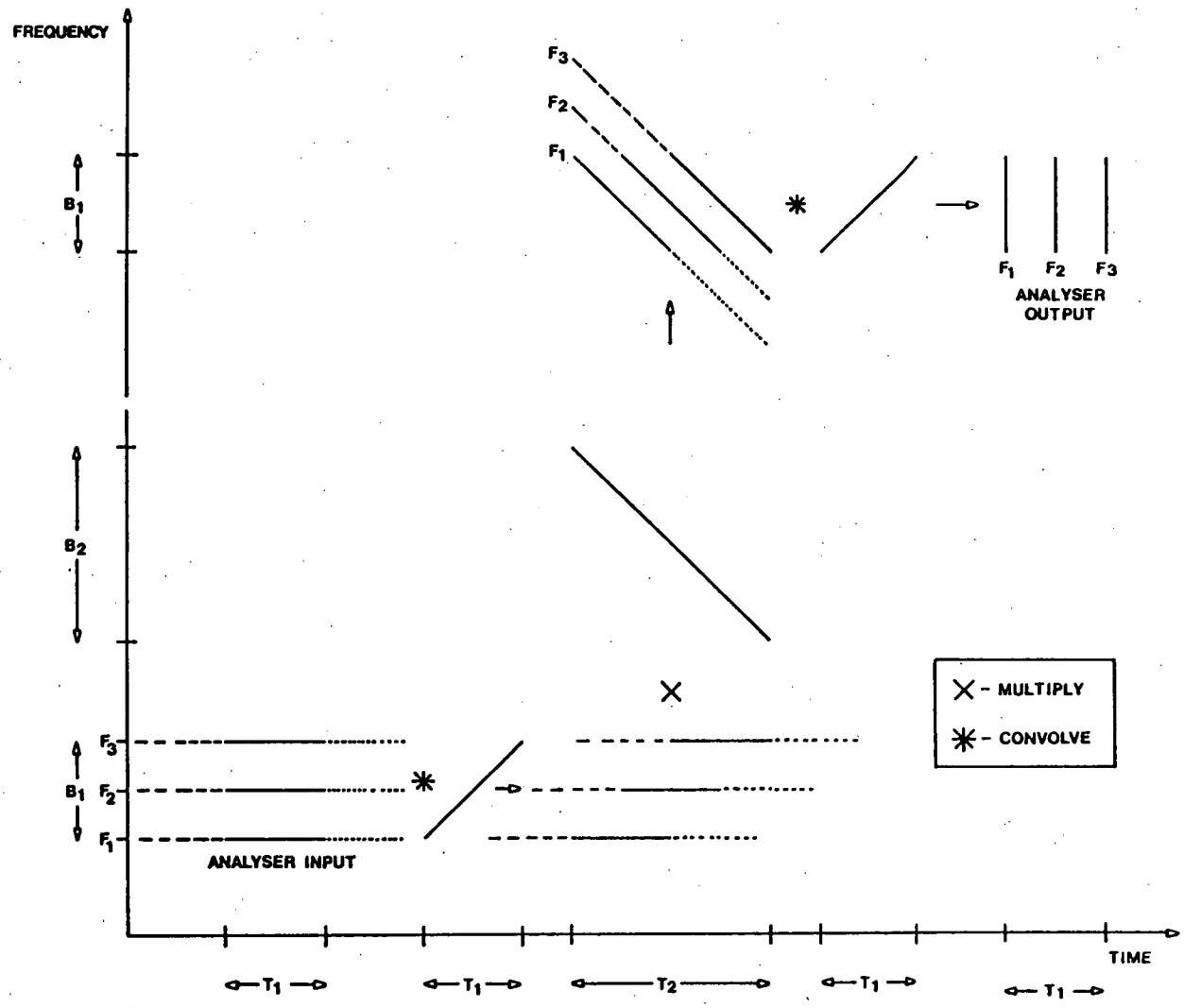


Fig.4-5: Frequency-time diagram showing the operation of the C-M-C transform algorithm.

time interval in which full compression occurs, the resolution function has a (sin x)/x form because the input sampling function is rectangular. The resolution function smears a line or CW input to a width of about $1/T_1$ where T_1, the duration of the analyser short chirp, is the duration of an input sample. Hence, the frequency resolution of a SAW analyser is about $1/T_1$. As in a radar compression loop (chapter three), compressed pulse sidelobe levels can be reduced by chirp weighting. However, the resulting increase in pulse width degrades the frequency resolution of the analyser.

For given chirp durations, frequency-time diagrams allow the duty cycle limitations and the response of a SAW processor to out-of-band signals to be predicted (refer again to figure 4-3(b)). To prevent confusion of in-band signals with out-of-band signals transformed in the previous or subsequent analyser cycle (a form of confusion called "aliasing" by analogy with the sampled data situation), it is necessary to ensure that only one input sample per long chirp duration is taken (Moule, 1982).

For an optimum chirp duration ratio ($T_2 = 2T_1$) the M-C-M analyser maximum duty cycle is 50%, so the input band is examined for only half the total time that the spectrometer is operating. In radioastronomy the corresponding sensitivity loss of $\sqrt{2}$ relative to a continuously operating analyser may be unacceptable, although the degradation factor is less than that for one-bit correlation spectrometers. Nearly continuous operation of the analyser (100% duty cycle) can be achieved if an input bandpass filter is used since corruption of valid data by out-of-band signals is then prevented. The shape factor of the input filter determines how closely the duty cycle can approach 100% before significant data corruption occurs.

The C-M-C Fourier processor would also be subject to a duty cycle limitation but for the fact that the first convolution stage (the pre-convolver) acts as a band limiter. SAW DDL's offer exceedingly good shape factors (Williamson et al., 1979) so the C-M-C duty cycle can be essentially 100%.

The M-C-M processor is theoretically the more satisfactory configuration for a SAW Fourier analyser because Fresnel effects inherent in the C-M-C approximation produce a higher far-sidelobe level on

compressed outputs. The M-C-M arrangement also has the advantage of requiring only two DDL's if a power spectrum analyser is required.

To appreciate the disadvantages of the M-C-M configuration the behaviour of the compressed pulse carrier component must be considered. In an M-C-M analyser the centre frequency of the compression operation is a function of the processor input frequency (figure 4-3(a)). Mixer chirps due to different input frequencies correlate in different physical regions of the convolver IDT (appendix one), precluding the incorporation of a weighting filter in the DDL. Chirp weighting in an M-C-M analyser is provided by either shaping the pre-multiplier chirp or by multiplying the input signal by a weighting function. Both approaches require a four-quadrant multiplier with wide input and output bandwidths. A suitable device is not available when bandwidths exceeding 20 MHz must be processed accurately (section 6.7).

In a C-M-C processor the convolution centre frequency remains constant because the convolver is filled by mixer chirps arising from inputs within the full-compression bandwidth (figure 4-5). Thus, in a C-M-C analyser the matched filter compression process (section 3.4) is approached and the SAW convolver can include a chirp weighting filter. Alternatively, a post-compression transversal network (figure 3-9) can be used.

Although inferior to an M-C-M processor in terms of sidelobe performance, the ease with which weighting can be incorporated into a C-M-C analyser is an important advantage in most applications. The stationary compression centre frequency also eases the response requirements of any post-transformation stages utilising the carrier as a reference for coherent processing. Further, in a C-M-C analyser the largest TB chirp is used only as an LO signal, so various passive and active techniques can be employed to generate the waveform from smaller TB chirps (Hall et al., 1982). Jack et al. (1980) point out that, for given SAW filters, the C-M-C configuration must be considered the optimum choice in terms of ultimate processor frequency resolution and bandwidth. They also note that for given parameters, the convolvers in a C-M-C analyser are of lower TB and hence lower insertion loss than the corresponding filter in the M-C-M arrangement. The C-M-C processor should therefore exhibit superior signal-to-noise ratio and dynamic range.

Davie and Arthur (1979) have examined both the M-C-M and the C-M-C Fourier analyser arrangements using computer simulation methods. They conclude that Fresnel effects inherent in the C-M-C configuration will produce a noticeable degradation in analyser performance, particularly in systems with TB<100. However, their study indicates that compression weighting can reduce many spurious sidelobes to an acceptable level. Thus, in applications where coherent post-transformation processing is needed and where optimum use of DDL TB is required, a weighted C-M-C analyser is the preferred type.

4.4 THE (C)-M-C ANALYSER

As mentioned, the final chirp can be omitted from an M-C-M analyser if power or amplitude spectrum analysis is all that is required. The resulting configuration, called an M-C-(M) arrangement by Moule (1982) and an M(short)-C(long) analyser by Jack et al. (1980), has the same strengths and weaknesses as a full M-C-M processor but offers a financial saving since only two DDL's are needed.

It is possible to produce a simplified version of a C-M-C analyser by discarding the first convolver (or pre-convolver). The resulting arrangement is known variously as a (C)-M-C (Moule, 1982), M(long)-C(short) (Jack et al., 1980), or compressive (or microscan) receiver spectrum analyser. It suffers from certain processing defects but offers nearly all the advantages of a full C-M-C analyser. Since it is the configuration chosen as the basis of instrumentation described in subsequent chapters, the characteristics of the (C)-M-C arrangement are discussed in detail in this section.

A block diagram of a typical compressive receiver is shown in figure 4-6 and a frequency-time diagram illustrating the spectrometer operation is given in figure 4-7(a). Provided pre-convolution signal chirps and the convolver impulse response chirp are opposite in slope, either up-chirp or down-chirp DDL's can be used as expansion and convolution stages, since the sign of the transform variable is normally of no significance. In some designs identical DDL's are used as expansion and convolution devices, chirp slope inversion then being performed in frequency translation stages. The analyser output appears at RF and for input signals in the full-compression band, the centre frequency of the RF spectrum is independent of the input frequency to the instrument.

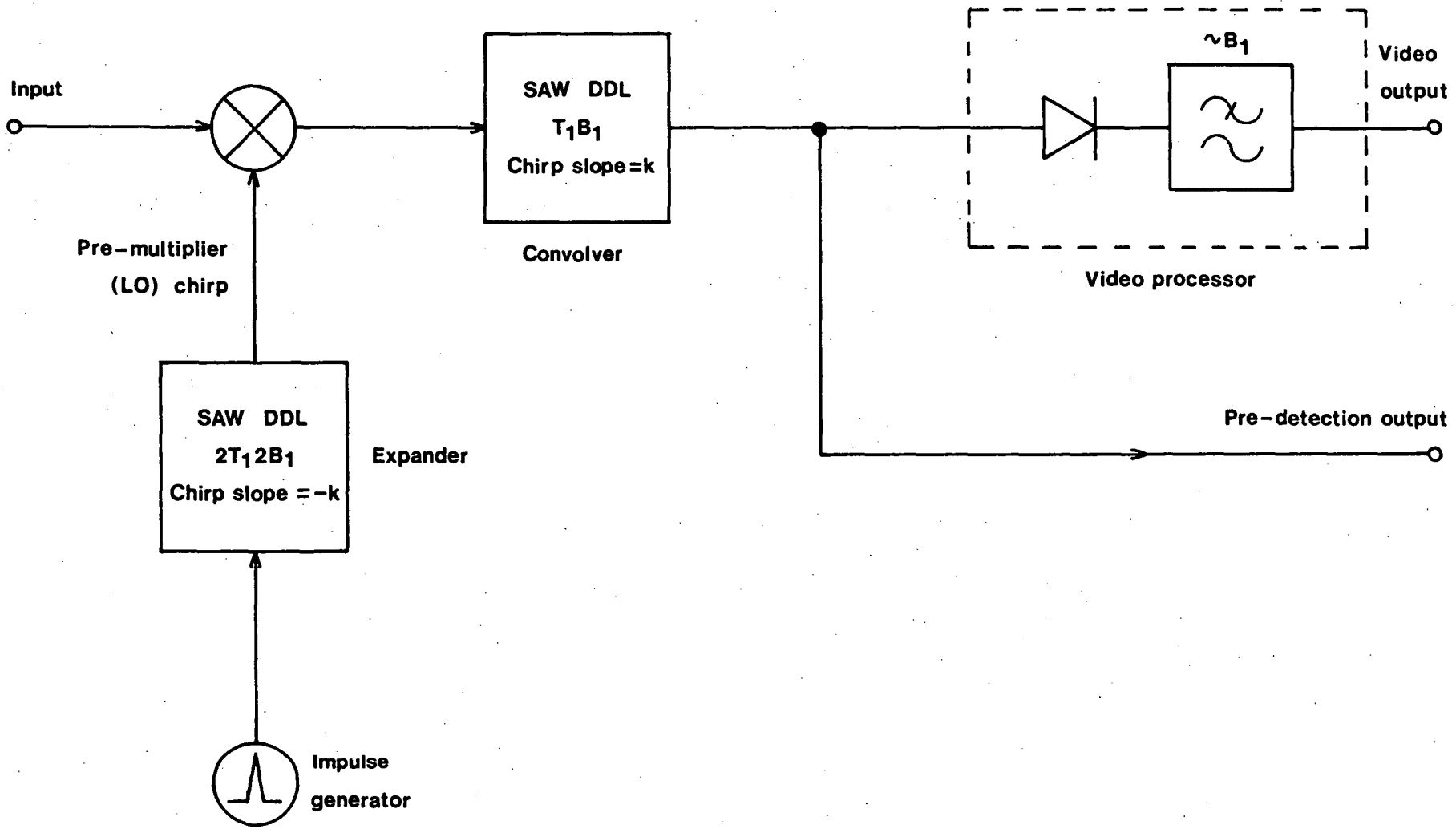


Fig.4-6: Block diagram of a basic SAW compressive receiver.

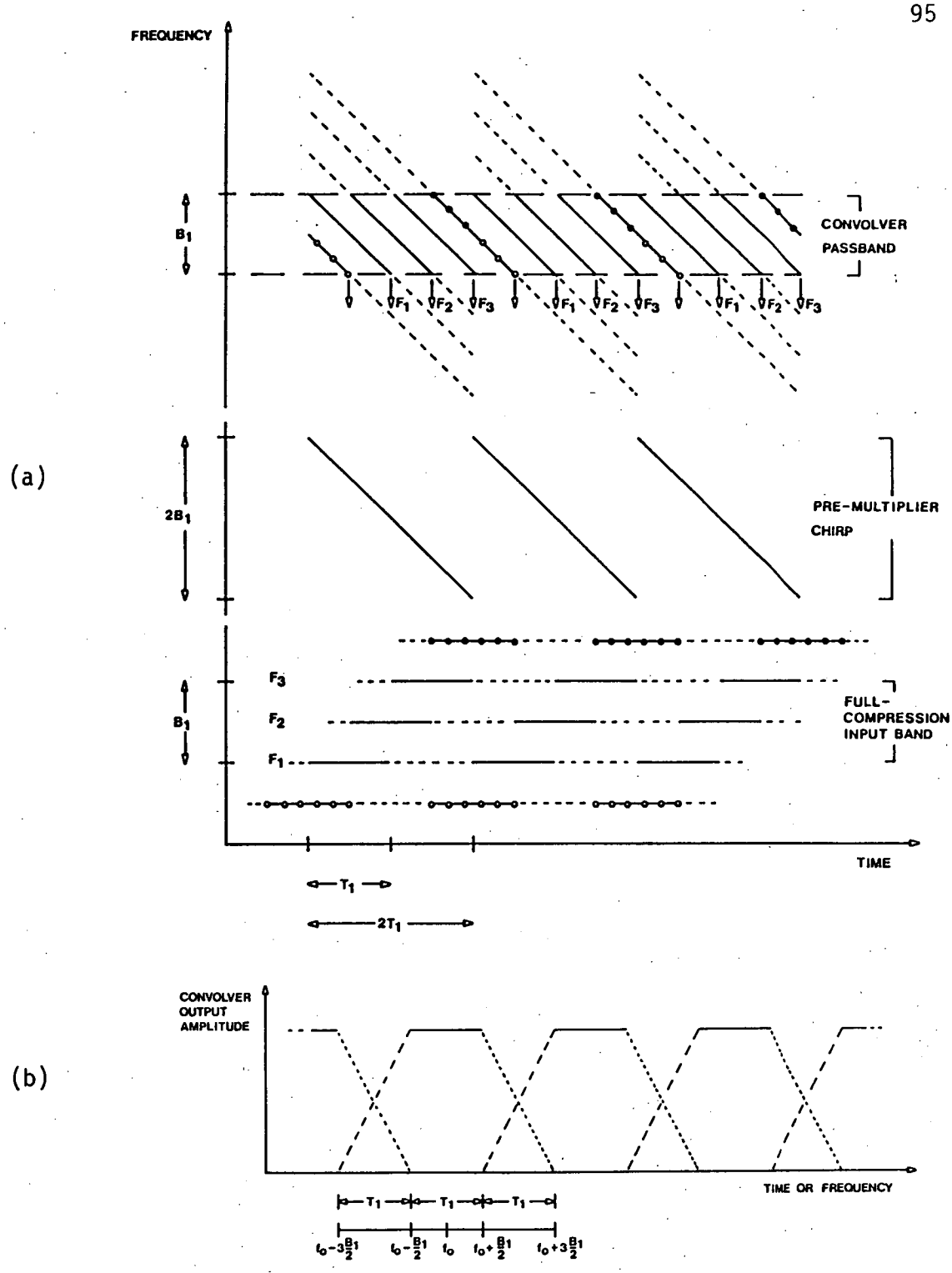


Fig.4-7: Operation of the compressive receiver shown in figure 4-6. In (a) a frequency-time diagram is given. Input signals are mixed with the down-chirp LO to produce down-chirp product signals. These signals are compressed in the up-chirp convolver. In this and subsequent frequency-time diagrams the absolute time delay of the convolver is neglected. The position in time of the emerging signals is then indicated by arrows. In (b) the idealized frequency response of the analyser is shown. The solid lines show the full-compression bandwidth in successive analyser operating cycles. It is clear that the duty cycle cannot exceed 50% if the full-compression data windows are to remain uncorrupted by out-of-band signals. The two out-of-band signals in (a) produce partially compressed outputs located at the intersection of the dashed and dotted lines in (b).

The 2:1 duration ratio of the pre-multiplier and convolver chirps ensures that signals at the edges of the input band are fully compressed. There is no pre-convolver to perform a band-limiting function so the maximum duty cycle is 50% if an input bandpass filter is not employed. The resultant receiver frequency response characteristic is illustrated in figure 4-7(b). In practical analysers the exact shape of the frequency response can be predicted by convolving the bandpass characteristic of signal chirps with the convolver frequency response. Figure 4-8 is a frequency-time diagram for a compressive receiver operating at 100% duty cycle and illustrates the need for a sharply band-limited input in this mode.

The disadvantage of the (C)-M-C arrangement is that pulsed input signals of shorter duration than the pre-multiplier chirp time ($T_2 = 2T_1$) are incorrectly transformed. As illustrated in figure 4-7(a), the input frequency-time space is not rectangular as it is in the full chirp transformer schemes but is trapezoidal. In effect, Fourier transformation takes place not on a time-gated sample of the input, f(t), but on a sample produced by sliding a gating function across f(t). It is this "sliding transform" which causes incorrect processing of short pulses.

The effect of the sliding transform on pulsed signals has been investigated by Otto (1976) and by Davie and Arthur (1979). The frequency resolution required to separate pulse harmonics is $1/T_p$, where T_p is the pulse duration. The processor resolution is about $1/T_1$, so if $T_1 < T_p < 2T_1$ the input pulse cannot be identified unambiguously. If $T_p < T_1$ a new processing regime is entered. The convolver output (amplitude and position) depends upon the pulse width, its input frequency and its phase relative to the operating cycle of the compressive receiver. Davie and Arthur discuss the effect in detail and offer experimental evidence for their conclusions.

SAW DDL chirp times of tens of micro-seconds or less are common, so longer duration signals can be considered to be CW or quasi-CW types for the purpose of spectrum analyser evaluation. In radioastronomy the sliding transform distortion is of no significance, except when observations of the pulsar micro-structure are undertaken.

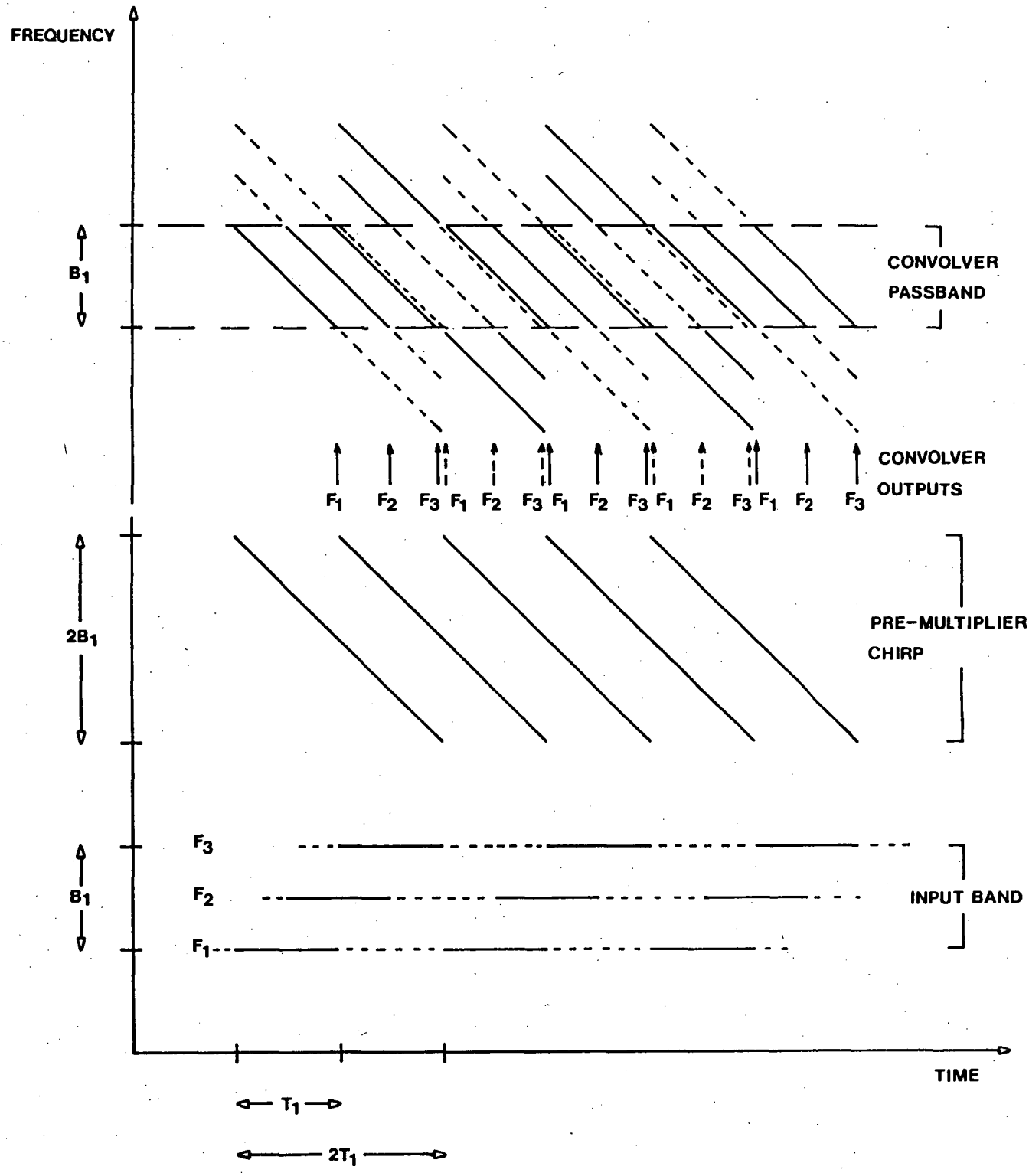


Fig.4-8: Frequency-time diagram for a compressive receiver operating at 100% duty cycle. In this mode the input band is sharply band-limited, allowing uncorrupted full-compression data windows from successive cycles to appear side-by-side at the convolver output.

The compressive receiver frequency-time diagram (figure 4-7(a)) leads to the inference that the compression process is identical to that in a radar situation (chapter three). During the development of prototype SAW analysers it was noted that the sidelobe levels of compressed pulses increased as the edges of the full-compression input passband were approached. This phenomonen, not previously reported, is explained by considering Fresnel ripple present on the LO chirp spectrum (figure 3-4) and not accounted for in a frequency-time diagram.

Band-edge signals must be mixed with the undulating extremeties of the pre-multiplier chirp to produce product chirps falling within the convolver passband. In practice, the pre-multiplier signal is not limited fully by the mixer and, in any case, Fresnel phase modulation is present. As predicted by paired echo theory, the modulation causes a higher sidelobe level for compressed pulses representing signals near the edges of the analysis bandwidth. The Fresnel modulation is mirrored at opposite ends of the pre-multiplier spectrum, so it is expected that these excess sidelobes will be generated preferentially on opposite sides of the main compressed pulse for input signals above and below centre frequency. Observation confirms the prediction and computer simulation (Arthur, 1983) produces similar results. The effect is more noticeable in low time-bandwidth product analysers where Fresnel ripple is most severe. For example, computer simulation of a compressive receiver with T_1B_1=55 and a Hamming weighted convolver shows excess sidelobe levels of 2 dB at the band edges.

A description of the major characteristics of SAW compressive receivers has been given by Hall et al. (1982). The receiver analyses a bandwidth B_1 in a time T_1 (figure 4-7) to give T_1B_1 spectral points. The resolution in frequency is approximately $\Delta f=B_1/T_1B_1=1/T_1$. Associated with each point is an amplitude and relative phase, the information being contained as the amplitude and phase of a stationary, multiplicative RF term at the convolver output.

The analyser can be likened to a filterbank having T_1B_1 channels of individual bandwidth $1/T_1$. In a filterbank the spectral information is available in parallel form; in a compressive receiver it emerges in serial analog form. The SAW analyser produces a continuous transform:

if a CW input is varied slowly the position in time of the compressive receiver output changes continuously, the -4dB width of the un-weighted compressed pulse being $1/B_1$ in time or $1/T_1$ in frequency. Typical chirp weighting schemes broaden the compressed pulse by a factor of 1.3. If a scanning spectrum analyser and a compressive receiver are designed to give the same frequency resolution, the compressive receiver yields a processing speed advantage of T_1B_1. For equivalent processing speeds the compressive receiver resolution advantage factor is $\sqrt{T_1B_1}$.

The analysis bandwidth of a compressive receiver is equal to the difference of the pre-multiplier chirp and the convolver chirp bandwidths. The output bandwidth is equal to the convolver bandwidth reduced by a compression weighting (or pulse broadening) factor where appropriate. For an un-weighted compressive receiver having an optimum chirp duration ratio the input and output bandwidths are equal.

The output signal-to-noise ratio defined for a single frequency resolution cell of a compressive receiver is the same as that for a filter matched to a CW input signal of duration T_1. Figure 4-7(a) shows that quasi-CW input signals falling within the nominal analysis bandwidth produce mixer chirps which correlate over the full time-scale of the convolver impulse response. Short pulse signals and out-of-band signals produce chirps which correlate partially, producing convolver outputs exhibiting lower compression gain and poorer sidelobe characteristics. Harrington and Nelson (1974) have examined compressive receiver sensitivity for various types of input pulse (see also the studies cited earlier in relation to sliding transform distortion).

The dynamic range specification of a SAW microscan receiver depends on the definition adopted. The single-signal against noise dynamic range often exceeds 60 dB but in a multi-signal environment sidelobes of strong signals dominate, limiting the dynamic range to around 35 dB. Furthermore, dynamic range and frequency resolution are related via the compressed pulse shape in a processing system where the receiver output is sampled. Moule (1982) estimates that the -35 dB resolution is typically three times worse than the -3 dB resolution.

4.5 PERFORMANCE OF SAW SPECTRUM ANALYSERS

In many applications, including radioastronomy, the need to produce high time resolution spectra makes SAW spectrum analysis attractive. The major disadvantage of SAW instruments, like other analog devices, is a lack of accuracy compared with digital spectrum analysers. Usually the speed advantage outweighs the accuracy limitation. Nevertheless, considerable commercial effort is devoted to improving the accuracy of SAW instruments by improving the response accuracy of dispersive delay lines.

The accuracy and speed performance of SAW spectrum analysers can be expressed in terms normally reserved for digital systems. Gautier (1981) points out that since dynamic range and accuracy are independent quantities in an analog system, it is necessary to define separately an equivalent number of bits for both quantities. If the single-signal against noise ratio is better than 60 dB the number of "dynamic range bits" exceeds ten (6 dB/bit).

The estimation of "accuracy bits" is more difficult. Gautier proceeds by considering uncertainties in the amplitude and phase spectra produced by a SAW analyser generating spurious sidelobes at a certain level. Typical chirp filters exhibit amplitude and phase ripples in a range of 0.1 dB and 1-2 degrees and, using paired echo theory (section 3.5), it can be shown that the resulting sidelobe level is around -35 dB. Hence, the equivalent accuracy of SAW systems is about six bits.

To compare the relative speeds of SAW and digital instruments it is possible to define an equivalent processing speed for a SAW spectrum analyser. Taking a radix-two digital FFT analyser as a reference, the number of floating point operations for an N-point transform is $5N.\log_2 N$ since the butterfly computational element (which performs typically ten basic operations) is invoked $(N/2).\log_2 N$ times (section 2.4; Gautier and Tournois, 1980; Gautier, 1981). In a SAW compressive receiver T_1B_1 points are produced in time T_1, giving an equivalent computation rate of $5B_1.\log_2(T_1B_1)$ floating point operations per second (FLOPS). For a modest SAW instrument having B=30 MHz and T_1B_1=64, the equivalent processing speed is around 1 GFLOPS, a factor of at least 50 faster than most digital instruments. Several modifications of the

standard FFT implementation are possible and high speed digital analysers are feasible. Nevertheless, with practical SAW systems offering equivalent computation rates exceeding 10 GFLOPS, current digital systems cannot compete on the basis of speed.

An important comparison which needs to be made between SAW and digital instrumentation is the stability and reproducibility of instruments using the respective technologies. Analog devices are more sensitive to environmental conditions than digital circuits and are usually more reliant on alignment for proper operation. In general, alignment needs to be repeated several times throughout the lifetime of the instrument.

The serial nature of a SAW spectrum analyser output means that spectrometer gain variations in both pre-transformation and post-transformation stages have an equal effect on all frequency cells. The common signal path is an advantage in complex instrumentation since systems using alternative analog technologies (such as filterbanks) may require extensive, real-time gain levelling techniques to be employed if cell-to-cell variations are to be minimised.

SAW spectrum analysers utilizing commercial DDL's are simple to construct, require little initial and no post-commissioning alignment, and can be made insensitive to environmental changes. Temperature is the environmental factor which has the greatest effect on SAW instruments. Fortunately, SAW DDL's are small enough to allow very effective temperature stabilization with conventional techniques.

Gautier (1981) shows that the time-bandwidth product of a SAW DDL remains constant with temperature, although changes in temperature produce changes in the absolute delay and the dispersive slope. In a SAW spectrum analyser the effect is to change the output time-frequency scaling factor and RF phase, assuming that the processor chirp signals remain matched in absolute slope. Chomiki (1984) gives the relationship between input frequency (f) and output time (t) as

$$t = T_{Ro} + T_{Co} - \frac{\alpha}{|k|}(F_{Ro} - F_{Co} - f), \qquad 4.5.1$$

where $k = B/T$ is the chirp slope in the cyclic frequency-time plane, and

T_{Ro} and T_{Co} are group delays at F_{Ro} and F_{Co}, the expander and convolver DDL centre frequencies. $F_{Ro} - F_{Co}$ is the analyser centre frequency (f_o), $\alpha = +1$ if the output delay increases with frequency and $\alpha = -1$ in the other case. Delays are measured from the instant of expander excitation.

The sensitivity of delay with temperature (θ) is

$$\frac{\Delta t}{\Delta \theta} = -\kappa\left(t + \frac{\alpha}{|k|} f\right) , \qquad 4.5.2$$

where κ is the substrate temperature coefficient ($-8.5 \times 10^{-5}/^{o}C$ for for YZ-$LiNbO_3$, $<3 \times 10^{-6}/^{o}C$ for ST-quartz). If $\alpha = -1$ and

$$f_o = F_{Ro} - F_{Co} = |k| \ (T_{Ro} + T_{Co}) \qquad 4.5.3$$

then $\Delta t/\Delta\theta = 0$ at f_o. As f changes, the maximum value of $\Delta t/\Delta\theta$ over the range $f_o \pm B_1/2$ is $\pm \kappa T_1$. The processor cell width is about $1/B_1$, so the maximum position error is $\kappa\Delta\theta T_1 B_1$ times the resolution. For $T_1 B_1 \lesssim 100$, the error remains less than a resolution cell for $\Delta\theta < 100^{o}C$ with an $LiNbO_3$ substrate.

If temperature changes cause the DDL responses in a SAW analyser alter differentially, the situation becomes analogous to a radar compression loop with a mismatch between expander and compressor chirp slopes (section 3.5). As in the radar case, the compressed pulse is degraded but chirp weighting reduces the distortion.

Gautier (1981) gives examples of analyser temperature stabilization and compensation schemes. More straightforward measures can also be used to minimise temperature effects. SAW devices should preferably be down-chirp to take advantage of the tendency towards self-compensation (appendix II) and should be bonded to a common metallic support to minimise differential response changes. For TB > 100, quartz DDL's are preferred if the added insertion loss is tolerable.

Although literature dealing with the mechanical aspects of SAW DDL's is not readily available, use of the devices in military airborne radar and spectral surveillance equipment indicates a high order of

reliability under the most adverse conditions. For practical purposes, digital and SAW spectrum analysers are assessed at roughly the same level of mechanical sensitivity.

PART B: POST-CONVOLVER PROCESSING

4.6 SPECTRUM ANALYSER OUTPUT DETECTION

The simplest output processing stage for a SAW analyser is an envelope detector followed by a lowpass filter (figure 4-9). If the detector diodes are operated in the small-signal forward conduction region an approximate square-law response can be obtained. Larger input signal levels drive the diodes into a linear detection mode and the output becomes proportional to the modulus of the input signal. The baseband, or video, signal resulting from the detection process has frequency components extending to the output bandwidth of the analyser. Since this bandwidth is of the order of the analysis bandwidth, the detector and associated filter must be designed in accordance with good high frequency engineering practice.

In practical detectors the diodes are hot-carrier (Schottky barrier) types capable of processing convolver output frequencies in excess of the 1 GHz upper limit on current SAW technology. High sensitivity Schottky diode detectors often exhibit more than 25 dB (referred to the input) square-law dynamic range. Hewlett-Packard, in a series of engineering application notes (AN907, AN923, AN963, AN969) give extensive design information for high sensitivity, wideband detectors which are suitable for processing SAW compressive receiver outputs.

Although diode detectors are normally used in wideband SAW analysers, more elaborate active detectors having a greater dynamic range have been described. Arsenault and Dolat (1981) describe a linear detector having a dynamic range of more than 32 dB. They use a highly limited (and phase compensated) replica of the convolver output as the reference for coherent detection in a mixer, which itself provides additional limiting. The convolver output resembles a double sideband supressed carrier signal with the form

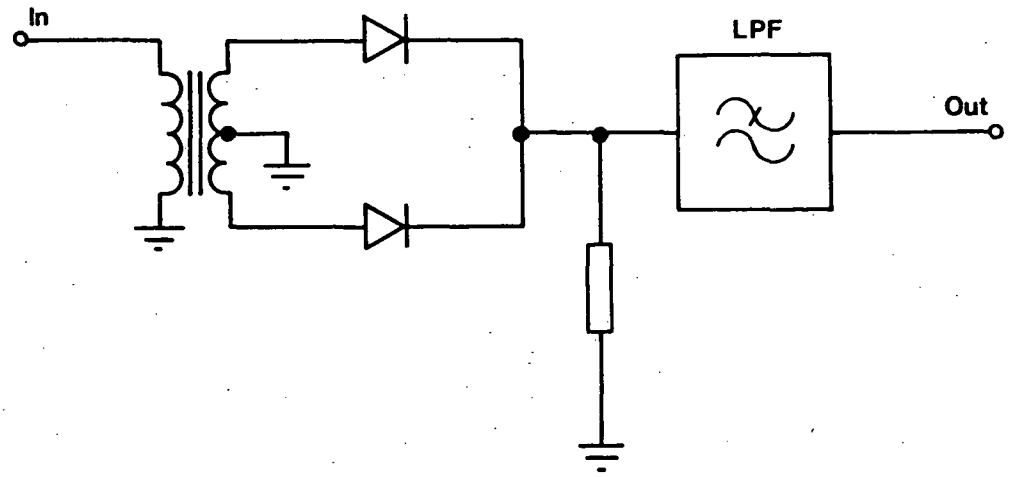


Fig.4-9: Full-wave detector circuit. The detector input comes from the convolver output in a compressive receiver or from the output of the second convolver in a C-M-C analyser. In a simplified M-C-M configuration the final multiplication is omitted and the convolver output detected.

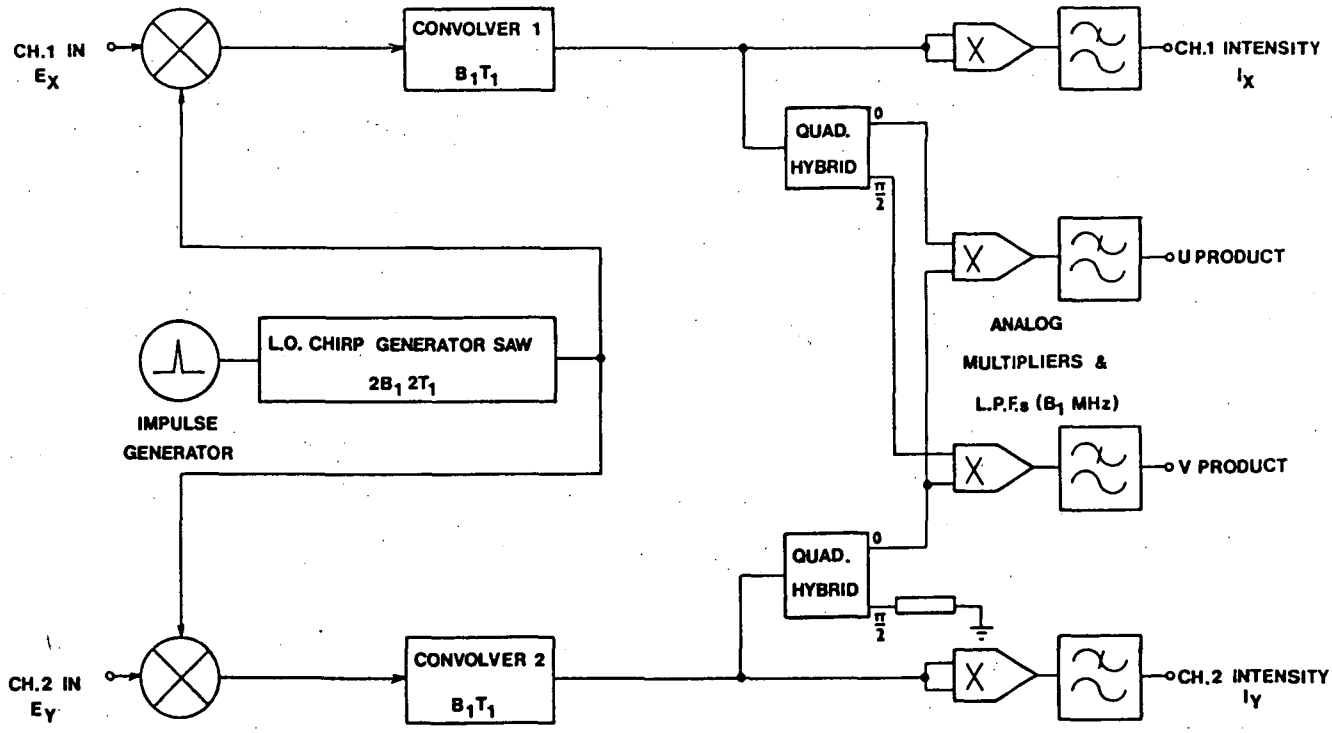


Fig.4-10: Basic form of a SAW polarimeter spectrometer using two compressive receivers.

$$g(t) = m(t)\cos \omega_0 t \quad ,$$

where m(t) is the spectral modulation. The output is small when the modulation is small, so severe clipping is necessary. Hall (1982) describes a multiplier or correlator which has been used as an accurate square-law detector in analysers with bandwidths of up to 30 MHz. This device is discussed in section 6.7.

The post-detection lowpass filter restricts the video bandwidth to the analyser output bandwidth, blocking high frequency components from the detector. These unwanted signals consist of RF components which may be centred at the convolver mid-frequency or its harmonics. The detector input is band-limited so the un-filtered output spectrum is zonal (band-like) in nature. The actual form of the spectrum depends on the detection law and upon whether the detector is a half-wave or full-wave type.

Butson et al. (1966) have examined the theoretical sensitivity of a compressive receiver as a function of post-detection bandwidth. They conclude that no one ratio of pre-detection to post-detection bandwidth is ideal, and they choose equal pre-detection and post-detection bandwidths to give a sensible compromise between compressed pulse shape degradation and reduction in sensitivity due to broadband noise in the video stages.

In applications where the output of a SAW analyser is sampled by a fast analog-digital converter, the presence of RF energy at the A/D converter input leads to aliasing problems. The post-detection filter must therefore have a steep enough roll-off in the frequency domain to prevent significant corruption of the sampled video signal by aliased RF components. However, a sharp frequency response gives a long impulse response. As the spectrum analyser output is in serial form, a long impulse response produces an effect similar to that seen in a filterbank analyser which has a high level of adjacent channel coupling. This is easily visualized if it is recalled that CW inputs produce outputs resembling time impulses. The video filter impulse response (or more exactly, the convolution of the impulse response and SAW analyser resolution function) therefore appears at the filter output. Thus, a long impulse response degrades the effective frequency resolution of the SAW analyser.

In practice, the post-detection filter must be a compromise between time and frequency domain requirements. If the ratio of the convolver centre frequency to bandwidth is large (small fractional bandwidth), video filters having a gentle frequency roll-off and non-oscillatory impulse responses can be used. Such filters include Gaussian and low-order Butterworth types. For large fractional bandwidths it is necessary to use sharper filters and in this case it is often beneficial to use video amplifiers with carefully designed roll-off characteristics as additional lowpass elements. The problems associated with a large fractional bandwidth are alleviated to some extent in a spectrometer using a full-wave detector, since the only fundamental convolver RF components at the detector output are then due to direct feed-through and to imperfections in the symmetry of the detector dynamic response (Middleton, 1960).

4.7 COHERENT SAW ANALYSERS

In both M-C-M and C-M-C Fourier analysers a complex spectrum is available at baseband if the final convolver output is multiplied either by quadrature chirp signals (in an M-C-M analyser) or quadrature CW signals (in a C-M-C analyser). This process recovers T_1B_1 complex spectral points and can be extended to the compressive receiver or (C)-M-C configuration to produce a complex sliding transform. It is assumed that lowpass filters which pass only baseband components are present at the baseband mixer outputs (Jack, 1978). In all cases the multiplications can be done using standard RF mixers since the LO signal is a uniform amplitude chirp or CW signal. Jack et al. (1980) have demonstrated the practical results obtained with such an arrangement.

Use of the multiplicative RF signal at a compressive receiver convolver output as a phase reference for more general coherent processing was mentioned by Harrington and Nelson (1974). The difficulty of making wideband, bi-linear, analog multipliers necessary in many potential applications has tended to discourage research into coherent SAW processors although two instruments have been described prior to the present study.

Gerard and Otto (1976), Gerard et al. (1977) and Estrick and Judd (1982) have described a radar coherent processor designed as a programmable pulse compression filter. They have used RF mixers as an approximation to fast, bi-linear multipliers and a dynamic range of less than 10 dB is apparent. Estrick and Judd claim that an intensive military research program has recently produced an analog multiplier with a usable input bandwidth in excess of 100 MHz. The output bandwidth is not stated but it must be greater than 60 MHz for correct operation of the radar processor.

Klose and Skudera (1981) give a description of a SAW interferometer direction finding instrument in which the RF outputs of two compressive receivers are combined in a microwave phase discriminator. In-phase and quadrature signals from the discriminator are detected to give a vector representation of the phase difference between the receiver inputs. If the receivers are connected to spatially separated antennas, an angle of arrival can be determined for each signal appearing at the output. The performance is limited by the dynamic range and linearity characteristics of the phase discriminator-detector combination but is adequate for most surveillance applications.

The SAW polarimeter spectrometer was conceived independently of the other instruments. The polarimeter, described by Hall et al. (1982), consists of two compressive receivers operated from a common pre-multiplier chirp (figure 4-10). The receiver outputs are square-law detected to yield intensity signals, from which quantities proportional to the Stokes parameters $I=I_x+I_y$ and $Q=I_x-I_y$ are derived in off-line processing stages. The analyser outputs are also passed through broadband quadrature hybrids and the in-phase and quadrature products formed in wideband, bi-linear analog multipliers developed for the application (Hall, 1982). The multiplier baseband outputs are directly proportional to the Stokes parameters U and V.

The similarity between the SAW polarimeter and the basic narrowband instrument illustrated in figure 1-2 is clear, although broadband RF splitting and phase shift elements are necessary in the SAW analyser to accommodate the entire output spectrum of the compressive receivers.

Since spectra emerge serially from each analyser, a single set of detectors and correlators can be used to form the required quantities. In effect, a spectrum for each Stokes parameter is available.

The underlying assumption in the polarimeter design is that the responses of the two SAW convolvers are well matched, so RF signals which are coherent at the processor inputs produce spectrum analyser outputs which are also coherent. Multiplication of the two analyser outputs then results in a quantity proportional to the vector product of the two output waveforms or, alternatively, an output proportional to the cross-spectrum of the input signals. SAW DDL fabrication methods give accurate convolver matching and the physical stability of the transversal filter realization (appendix one) guarantees that coherence is maintained between the two processor RF channels.

In a coherent processor based on compressive receivers, mixer chirps resulting from in-band inputs fill the SAW convolvers (figure 4-7(a)). The pre-multiplier chirp splitting network, signal mixers or amplifiers may produce a non-zero differential phase response between the two spectrum analysers but, disregarding Fresnel effects, the compression characteristics of either analyser do not alter with input frequency. Thus, if the processor performs correctly for signals at centre frequency, the overall performance will be satisfactory, provided the phase shifts mentioned can be minimised.

The considerations mentioned in section 4.6 relating to output processing also apply to a polarimeter or similar coherent SAW processor. In addition, two points applying specifically to coherent instruments can be made. Firstly, unwanted pre-correlation differential phase shift must be minimized, so identical RF components and cable lengths must be used in the two spectrum analyser signal paths. Although a broadband quadrature phase shift is required in only one output, the use of two quadrature hybrids in the arrangement shown in figure 4-10 compensates for the hybrid insertion phase. Secondly, care must be taken to minimise differential group delays in networks following the detectors and correlators. Differential delays result in time (and therefore frequency) skew between the four output channels of the instrument.

4.8 DIGITAL PROCESSING METHODS FOR SAW ANALYSER OUTPUTS

In many applications SAW spectrum analysers require only such basic post-detection processing equipment as a CRT display and associated timebase trigger circuitry. In some applications, including radioastronomy, further digital processing is necessary. In all such processing, either sensitivity or time resolution must be sacrificed. The sacrifice is necessary because of the very high output data rate of SAW analysers. Assuming one sample per resolution cell and an analyser output bandwidth of B_0 ($\sim B_1$), the data rate is B_0 words/second. Even with analysers of moderate bandwidth, this rate exceeds the sampling and processing capabilities of most digital instruments. For example, a typical computer may be capable of transferring 1-2 million words/second in a fast direct memory access mode. A single 20 MHz SAW compressive receiver requires a 20 million word/second transfer. The real situation is worse because the computer must process the transferred data. Even the elementary operation of continuous spectral integration is therefore beyond the capabilities of present-day computers.

In an effort to overcome the data rate problem, fast dedicated digital processors have recently been implemented. Moule (1982) describes devices which quantize wideband (200 MHz) compressive receiver outputs to one-bit resolution and store the results for many sweeps in high speed data buffers able to be read by a slow processor. Such elementary processing indicates only the presence or absence of a signal in a given frequency cell and cannot be regarded as a general technique. Alldritt et al. (1978) outline a spectrum integrator similar in concept to the unit described in chapter eight but having a maximum sampling frequency of 10 MHz. This limit, as well as the physical complexity of their approach, makes their integrator unattractive for applications requiring a large bandwidth or the use of several integrators.

These solutions exemplify two basic approaches to the data rate problem. Both involve the introduction of buffer hardware between a high speed A/D converter at the SAW analyser output and a slower digital processor. In the first method of data acquisition the SAW spectrometer is triggered, its output sampled and the data written to a high speed buffer.

The slow processor reads the buffer, processes the data and then triggers the SAW analyser to recommence the cycle. The operation is wasteful of sensitivity since the duty cycle of the spectrometer is very small, perhaps less than one percent.

The second approach maintains sensitivity but degrades the time resolution. In this method, which is the appropriate one for radioastronomy, a high speed integrator is interposed between the A/D converter and the slow processor (figure 4-11). In the unit described in chapter eight, successive analyser scans are added synchronously and when the integration is complete the integrated spectrum is written to a data buffer. While the next integration is proceeding the slow processor reads the buffer, retrieves the integrated spectrum and either processes the data or transfers it to a mass storage device. The important point is that the SAW analyser runs at its normal duty cycle and is unaffected by speed limitations in the data processing system. In a spectrometer utilising a spectrum integrator the minimum integration time, or best time resolution, is set by the speed at which the slow processor can accept data.

4.9 SAMPLING SAW SPECTRUM ANALYSER VIDEO OUTPUTS

Envelope detection of a SAW analyser output results in a waveform having a baseband (video) spectrum extending to B_0, the analyser output bandwidth. If the waveform is treated as a general signal the minimum sample frequency is $2B_0$, corresponding to Nyquist sampling. However, chirp transformation enforces an ordering between input frequency and output time so the waveform is not general in the usual sense. The position of a particular frequency resolution cell is determined precisely on the time waveform, so very low sample frequencies may yield the required spectral information in some circumstances.

The spectrum of compressed pulses emerging from a SAW analyser is truncated by virtue of the finite analyser chirps, so a sample frequency of $2B_1$ (B_1 is the short chirp bandwidth) is certainly acceptable. In a compressive receiver this criterion corresponds to sampling at the Nyquist rate for the unweighted convolver. When a weighted convolver is used (resulting in broader compressed pulses and

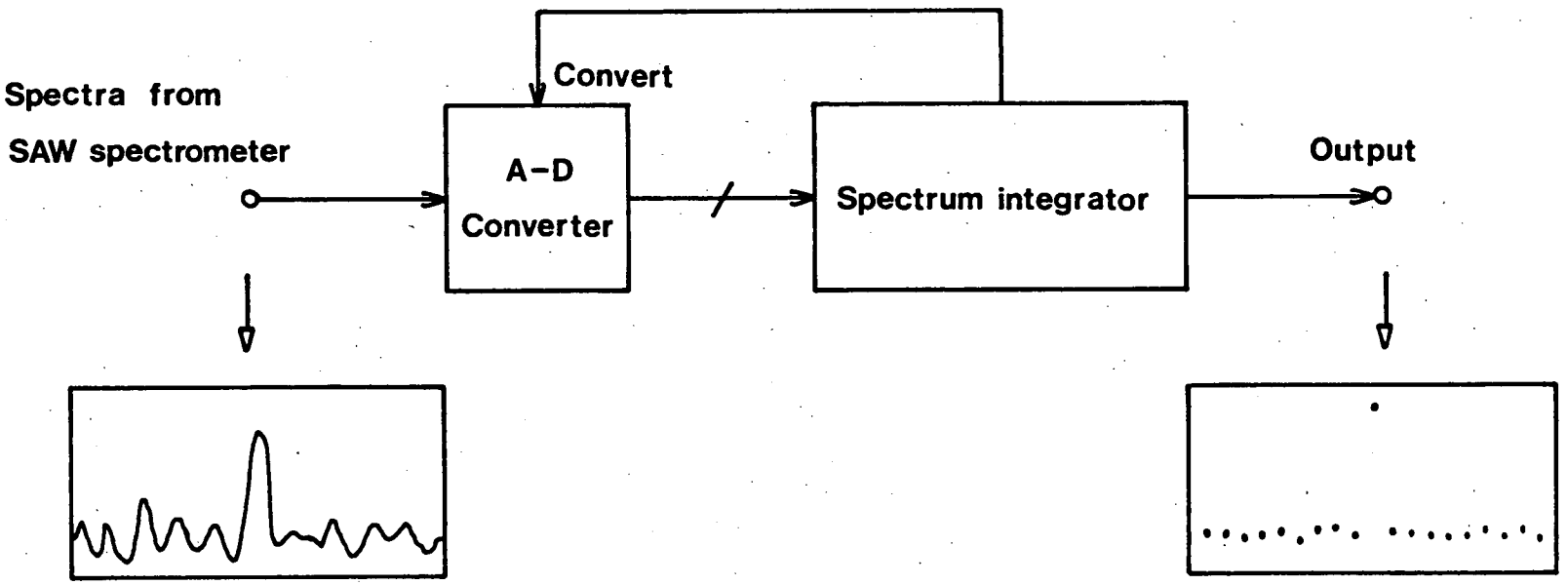


Fig.4-11: Operation of a digital spectrum integrator. Output spectra from the SAW spectrum analyser are sampled and added synchronously to produce an integrated spectrum.

a smaller output bandwidth), the convolver frequency response may be so far down at the passband edges that a lower sample frequency will suffice.

The choice of sampling frequency depends on the application envisaged for the SAW spectrometer. In many cases it is permissible to treat the analyser as a bank of frequency-contiguous receivers and to formulate sampling requirements using this analogy. For example, if an estimate of spectral power at only two spot frequencies is required, only two samples of the output waveform need be taken. If integrated spectral intensity over a certain frequency interval is needed it may be appropriate to restrict deliberately the frequency resolution of the analyser prior to sampling, perhaps by the introduction of an additional post-detector lowpass network.

As a rule, sampling the spectrum analyser at B_o produces around TB substantially independent samples. The resulting spectral sampling characteristic is similar to that obtained with a filterbank spectrometer having adjacent filters overlapping at the -3 dB points. Note that TB is the effective time-bandwidth product of the analyser; in an unweighted instrument $TB = T_1B_1$. A sample frequency of B_o is therefore the minimum frequency at which full use is made of the analyser time-bandwidth product. However, sampling faster than B_o can produce worthwhile results because the shape of the frequency resolution cell is not rectangular but is more like a conventional Gaussian filter. Over-sampling therefore tends to fill in spaces between the resolution cell response maxima which are obtained when a sample frequency of B_o is used. Sampling faster than $2B_o$ produces adjacent samples which are highly correlated and so the extra samples contribute little additional information in any reconstruction of the spectrum.

These considerations are particularly relevant to wideband SAW analysers where it may be impossible to acquire data with a clock frequency of B_o, let alone $2B_o$. In such cases the options are either to use interleaved digital processors or to sample the spectrometer output at as high a frequency as possible given technical and economic constraints. In the second approach potential ambiguities exist if

attempts are made to reconstruct spectra having features which vary over a scale of less than twice the interval between spectral samples.

In many applications the ambiguity need not in itself be a deterrent to the use of an otherwise suitable SAW instrument. Even if the sample interval is large compared with the scale of spectral variation in the radiation being observed, the analyser may be useful for some measurements. In the study of pulsars for example, useful spectral measurements are made with narrowband receivers tuned to frequencies separated by tens or hundreds of megahertz even though scintillation effects are known to produce short term intensity variations on a scale of hundreds of kilohertz.

CHAPTER 4 - REFERENCES

Alldritt, M., Jones, R., Oliver, C.J. and Vaughan, J.M.: The processing of digital signals by a surface acoustic wave spectrum analyser. J. Phys. E: Sci. Instrum., Vol.11, 1978, pp.116-119.

Arsenault, D.R. and Das, P.: SAW Fresnel transform devices and their applications. Proc. 1977 IEEE Ultrasonics Symp., pp.969-973.

Arsenault, D.R. and Dolat, V.S.: Compact multiple channel SAW sliding window spectrum analyser. Proc. 1981 IEEE Ultrasonics Symp., pp.220-225.

Arthur, J.W.: Private communication, 1983.

Butson, P.C., Lucas, W.J. and Thompson, G.T.: Theoretical assessment of the use of pulse compression in a panoramic receiver. Proc. IEE, Vol.113, No.5, May 1966, pp.725-739.

Chomiki, M.: How to reduce the effects of temperature variations in SAW chirp transform processors. To be published.

Cook, C.E.: A general matched filter analysis of linear FM pulse compression. Proc. IRE, Vol.49, p.831, 1961.

Das, P. and Shklarsky, D.: SAW implemented real-time Hilbert transform and its application in SSB. Proc. 1979 IEEE Ultrasonics Symp., pp.752-756.

Davie, M.C. and Arthur, J.W.: A comparative study of SAW compressive receivers for radar intercept. Racal-MESL Microwave Ltd., Internal report, Edinburgh, 1979.

Estrick, V.H. and Judd, G.W.: A SAW analog correlator using the chirp transform. Microwave Journal, Vol.25, No.7, July 1982.

Gerard, H.M. and Otto, O.W.: A programmable pulse compression filter utilizing chirp transform correlation. Proc. 1976 IEEE Ultrasonics Symp., pp.371-375.

Gerard, H.M., Yao, P.S. and Otto, O.W.: Performance of a programmable radar pulse compression filter based on a chirp transformation with RAC filters. Proc. 1977 IEEE Ultrasonics Symp., pp.947-951.

Gautier, H.: Digital processors using SAW devices. Proc. 1981 IEEE Ultrasonics Symp., pp.206-219.

Gautier, H. and Tournois, P.: Signal processing using surface acoustic wave and digital components. Proc. IEE, Vol.127, Pt.F, No.2, April 1980, pp.92-98.

Hall, P.J.: A multiplier/detector for SAW spectrum analysers. Electron. Lett., Vol.18, No.16, Aug. 1982, pp.699-700.

Hall, P.J., Hamilton, P.A. and McCulloch, P.M.: A correlation Stokes polarimeter using a dual channel SAW Fourier analyser. JEEE (Aust.), Vol.2, No.3, Sept. 1982, pp.150-155.

Harrington, J.B. and Nelson, R.B.: Compressive receiver uses SAW devices. Microwave Journal, Vol.17, No.9, Sept.1974, pp.57-62.

Jack, M.A.: Ph.D. thesis. University of Edinburgh, 1978.

Jack, M.A., Grant, P.M. and Collins, J.H.: The theory, design and applications of surface acoustic wave Fourier transform processors. Proc. IEEE, Vol.68, No.4, 1980, pp.450-468.

Klose, D.R. and Skudera, W.: A SAW interferometer direction finding and frequency identification method. Proc. 1981 IEEE Ultrasonics Symp., pp.392-394.

Middleton, D.: "An introduction to statistical communication theory". Ch.13.4, McGraw-Hill, 1960.

Morgan, C.: M.Sc. thesis. University of Manchester, 1982.

Moule, G.L.: SAW compressive receivers for radar intercept. Proc. IEE, Vol.129, Pt.F, No.3, June 1982, pp.180-186.

Nichols, B.: Transform receivers for ECM applications. Microwave Journal, Vol.25, No.10, October 1982, pp.113-127.

Oliver, C.J.: The measurement of optical spectra using a surface acoustic wave spectrum analyser. J. Phys. D: Appl. Phys., Vol.11, 1978, pp.2499-2507.

Oliver, C.J.: The statistical properties of surface acoustic wave spectrum analysers used in the photon counting mode. J. Phys. D: Appl. Phys., Vol.13, 1980, pp.1577-1588.

Otto, O.W.: The chirp transform signal processor. Proc. 1976 IEEE Ultrasonics Symp., pp.365-370.

Roberts, J.B.G., Moule, G.L. and Parry, G.: Design and application of real-time spectrum analyser systems. Proc. IEE, Vol.127, Pt.F, No.2, April 1980, pp.76-91.

Williamson, R.C., Dolat, V.S., Rhodes, R.R. and Boroson, D.M.: A satellite-borne SAW chirp transform system for uplink demodulation of FSK communication signals. Proc. 1979 IEEE Ultrasonics Symp., pp.741-747.

CHAPTER 5 QUANTIZATION OF SAW PROCESSOR OUTPUTS

5.1 INTRODUCTION

In many applications where spectra with moderate to high signal-to-noise ratios are processed it is reasonable to assume that a SAW spectrum analyser output ought to be quantized to a resolution of five or six bits since this is the approximate accuracy of the analog Fourier transformer (section 4.5). In radioastronomy spectra with very low signal-to-noise ratios must be digitized and integrated. This chapter investigates the characteristics of a SAW radiometer (spectrometer and integrator) with quantizers of various resolutions. The performance is described in terms of signal-to-noise ratio (SNR) and dynamic range in any one frequency resolution cell at the integrator output. Digital integration of analog correlator outputs is considered briefly. Sections 5.1 and 5.2 contain review material but the discussion of radiometer performance in subsequent sections is original.

The frequency resolution (Δf) of an un-weighted SAW analyser is $1/T_1$, so the output bandwidth-time product (not to be confused with the time-bandwidth product of chirp signals within the analyser) in a given resolution cell is unity. At 100% duty cycle a SAW analyser produces an independent sample from each resolution cell every $1/\Delta f$ seconds, the same rate as a conventional IF stage of bandwidth Δf. A particular cell therefore resembles a narrowband receiver having no post-detection smoothing. This analogy is pursued in the remainder of the chapter.

Consider initially the case where a CW signal with high SNR is detected using a linear detector at the spectrum analyser output. If the detector output is quantized and integrated the maximum dynamic range at the integrator output is about 6 dB per quantizer bit. For square-law detection the dynamic range referred to the spectrum analyser input is reduced to 3 dB/bit. If sufficient quantizer resolution is maintained to ensure that quantization noise is very small, the SNR after the addition of N independent samples is given by

$$(SNR)_{(out)} = (SNR)_{(in)} \sqrt{N} , \qquad 5.1.1$$

where $(SNR)_{(in)}$ is the SNR at the quantizer input.

To predict the performance of the radiometer with noisy input signals, the form of the signal probability density function (p.d.f.) at the quantizer input must be known. In this chapter the signal input to the spectrometer is assumed to be white noise (Gaussian amplitude statistics), modelling accurately the usual astronomical situation. The convolver output in a given cell is also Gaussian in the sense that successive time samples of the spectral amplitude from this cell are drawn from a Gaussian p.d.f.

In a coherent detection system the convolver output is translated to baseband using the quadrature mixing arrangement described in section 4.7. If the baseband signals are processed digitally, the signal in each cell at the input of both quantizers has Gaussian statistics. The quantization requirements for such signals are well known from the theory of digital correlators and are summarised by Cooper (1976).

If the SAW analyser output is envelope detected prior to quantization the situation is complicated because the signal at the quantizer input is no longer Gaussian. In the next section the theory of detection and integration in a conventional receiver having an envelope detector followed by an analog integrator is reviewed. In section 5.3 the receiver performance with analog and digital integrators is compared. The results are applied directly to SAW analysers using the analogy between a single resolution cell and a narrowband receiver.

5.2 DETECTION AND INTEGRATION OF NARROWBAND NOISE

The usual single channel receiver configuration is shown in figure 5-1. Full-wave envelope detection is assumed with the detector transfer function being either linear or quadratic in form. The post-detection lowpass filter excludes high frequency components and may pass all or part of the baseband spectrum. The baseband spectrum extends to Δf and if $f_c < \Delta f$ the filter acts as an integrator; if $f_c \ll \Delta f$ many independent noise pulses from the detector are summed. By the central limit theorem the noise fluctuation about the mean integrator output is then essentially Gaussian, regardless of the p.d.f. at the detector output.

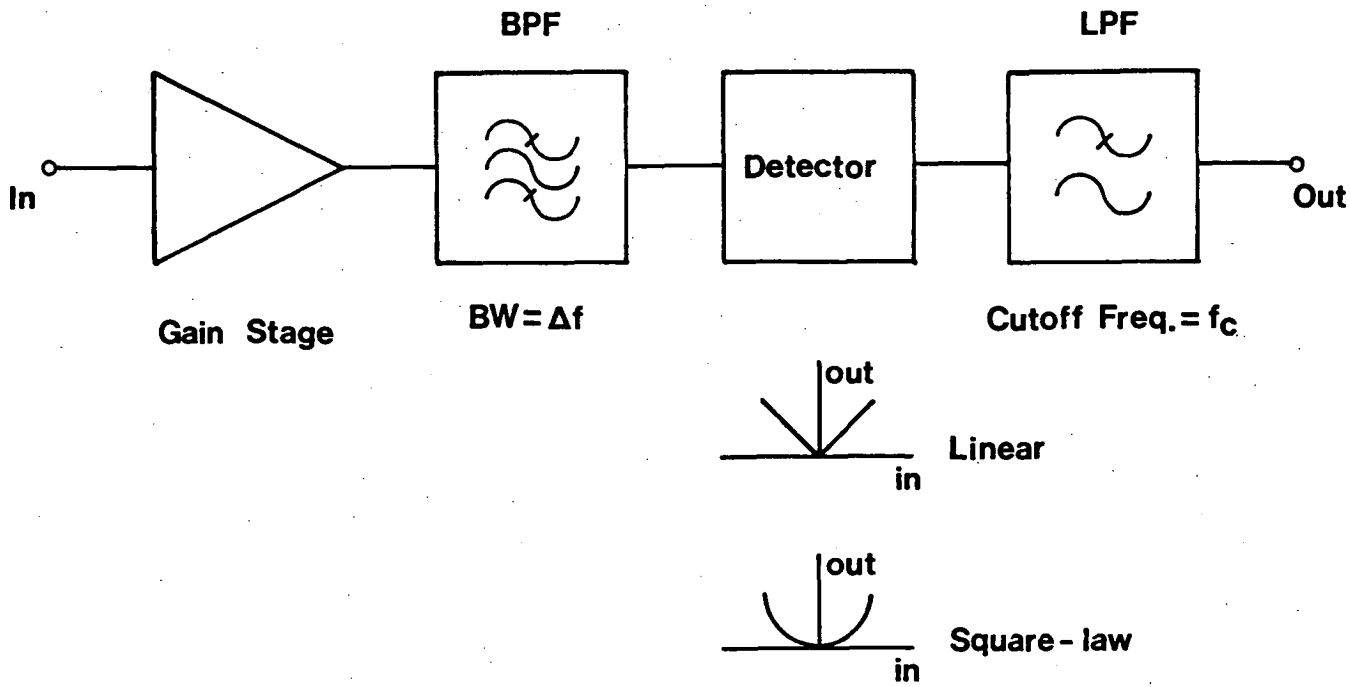


Fig.5-1: Narrowband receiver with full-wave envelope detector.

In a SAW spectrum analyser the baseband spectrum is intact, so quantizing the video output is analogous to quantizing the output of the receiver in figure 5-1 with $f_c = \Delta f$. There is no post-detection integration prior to quantization so the form of the p.d.f. at the quantizer input is not Gaussian but rather a Gaussian distribution transformed according to the transfer function of the detector. In subsequent analyses it is assumed that the post-detection filter (figure 5-1) has an ideal lowpass response with infinite attenuation above f_c and a flat passband below f_c.

The pre-detection signal is zero-mean Gaussian, the p.d.f. being

$$p(b) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left[-b^2/(2\sigma^2)\right] \quad , \qquad 5.2.1$$

where σ is the r.m.s. level. The channel is band-limited so a narrowband noise representation can be used (Middleton, 1960). Using this representation, the random signal can be written as

$$b(t) = V(t) \cos\left[\omega_0 t + \phi(t)\right] \quad , \qquad 5.2.2$$

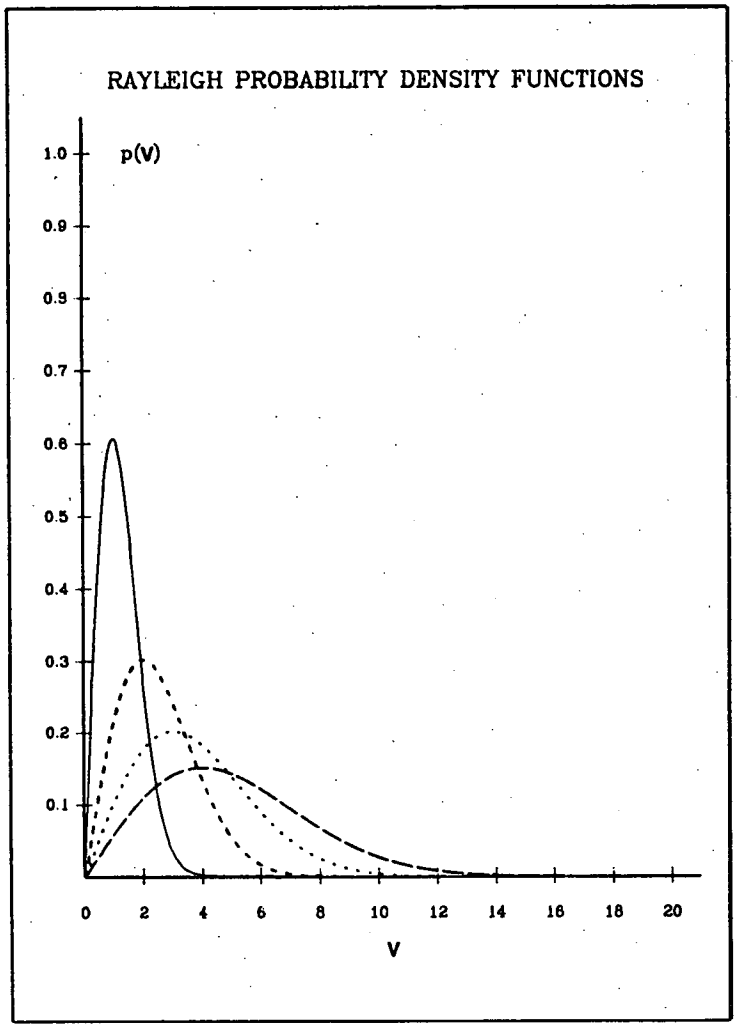
where $V(t)$ and $\phi(t)$ are envelope and phase terms which vary slowly compared with $\cos\omega_0 t$, a centre frequency variation. The phase is uniformly distributed over a range $0\text{-}2\pi$ radians and the envelope is Rayleigh distributed with a p.d.f. given by

$$\begin{aligned} p(V) &= (V/\sigma^2) \exp\left[-V^2/(2\sigma^2)\right] \qquad V \geq 0 \\ &= 0 \qquad\qquad\qquad\qquad\qquad\quad V < 0. \end{aligned} \qquad 5.2.3$$

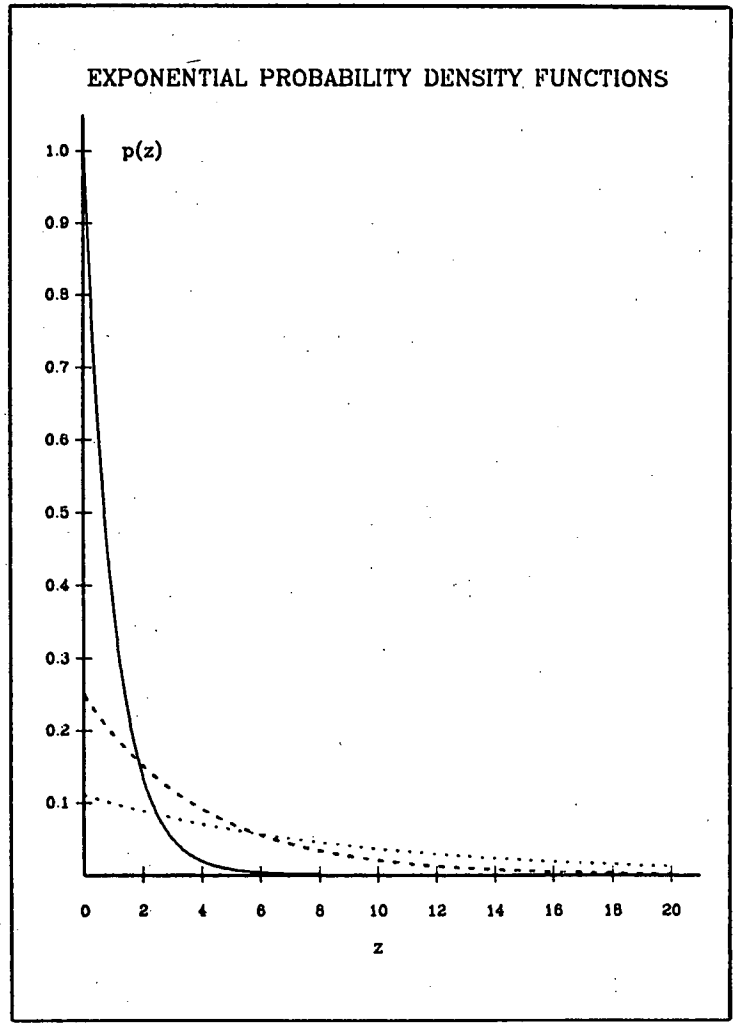
If $b(t)$ is detected using a full-wave linear (or modulus) detector, $p(V)$ is the p.d.f. at the quantizer input. Figure 5-2(a) shows the form of the distribution for several values of σ.

If $b(t)$ is square-law detected, the video filter output is $V^2(t)/2$. Substituting for the square of the envelope in equation 5.2.3 and conserving transformed probability ($p(V)dV = p(z)dz$) gives

$$\begin{aligned} p(z) &= (1/\sigma^2) \exp(-z/\sigma^2) \qquad z \geq 0 \\ &= 0 \qquad\qquad\qquad\qquad\quad z < 0, \end{aligned} \qquad 5.2.4$$



(a)



(b)

Fig.5-2: Probability density functions. Rayleigh distributions are shown in (a) while exponential distributions are shown in (b). The solid, light dashed and dotted lines are for σ = 1,2 and 3 respectively. The heavy dashed line in (a) is for σ = 4.

assuming the square-law device to have a transfer curve $y=b^2$ (Davenport and Root, 1958). This p.d.f. is plotted in figure 5-2(b) and is an exponential distribution. An alternative and equally valid view is that p(z) is a χ^2 (chi-squared) distribution having two degrees of freedom. Some important properties of Rayleigh and exponential distributions are summarised in table 5-1.

Table 5-1 Properties of Rayleigh and Exponential P.D.F.'s

Distribution	Mean	Variance	Standard Deviation
Rayleigh	$\sigma\sqrt{\frac{\pi}{2}}$	$\sigma^2[2-\frac{\pi}{2}]\sim 0.43\sigma^2$	$\sigma\sqrt{2-\frac{\pi}{2}}\sim 0.66\sigma$
Exponential	σ^2	σ^4	σ^2

Note that the mean and standard deviations are equal for the exponential distribution.

A full analysis of post-detection integration following a square-law detector is surprisingly difficult. Kac and Siegert (1947), Emerson (1953) and Meyer and Middleton (1954) have attacked the problem using various methods. A simplified approach is to consider the integrator as a device which sums N independent χ^2 variates, each having two degrees of freedom. The sum is also χ^2 distributed and has 2N degrees of freedom. For large N, the resulting p.d.f. approaches a normal distribution having a non-zero mean (d.c. offset). In fact, for N~10 the p.d.f. begins to approach Gaussian form (Whalen, 1971).

If a linear detector is used, the problem is intractable in the strict analytical sense because the sum of Rayleigh distributed random variables is not known in closed form. Robertson (1968) has used a type-A Gram-Charlier (G-C) series to approximate the integrator output p.d.f. over a given range. The G-C series is an expansion of the actual p.d.f. in terms of a Gaussian function and its derivatives.

When the output of either detector type is quantized prior to integration a full analysis becomes exceedingly difficult because the p.d.f. of the signal to be integrated is modified by the

presence of quantization noise. In the following investigations numerical simulation techniques have been used in an attempt to provide insight into the performance of systems having a post-detector quantizer.

5.3 PERFORMANCE OF DIGITAL POST-DETECTION INTEGRATORS

Two aspects of digital integrator performance have been examined using simulation methods. Initially, the linearity and SNR performance with a given quantizer resolution was examined as a function of quantizer input level; for the purpose of comparison it was assumed that the number of post-detection integrations (N) was 1000 (a realistic figure) and that the full-scale quantizer input was 1.0 V. The second study examined the relative SNR performance of channels having linear and square-law detectors for different values of N. Throughout the linearity and SNR investigations a simulated un-quantized detector output was retained to provide both a reference standard and a check on scaling and computational methods. This is referred to as an "analog" system. In the second study only an analog integrator was simulated since operation of a quantized system in its optimum input region ensures that the analog performance is approached.

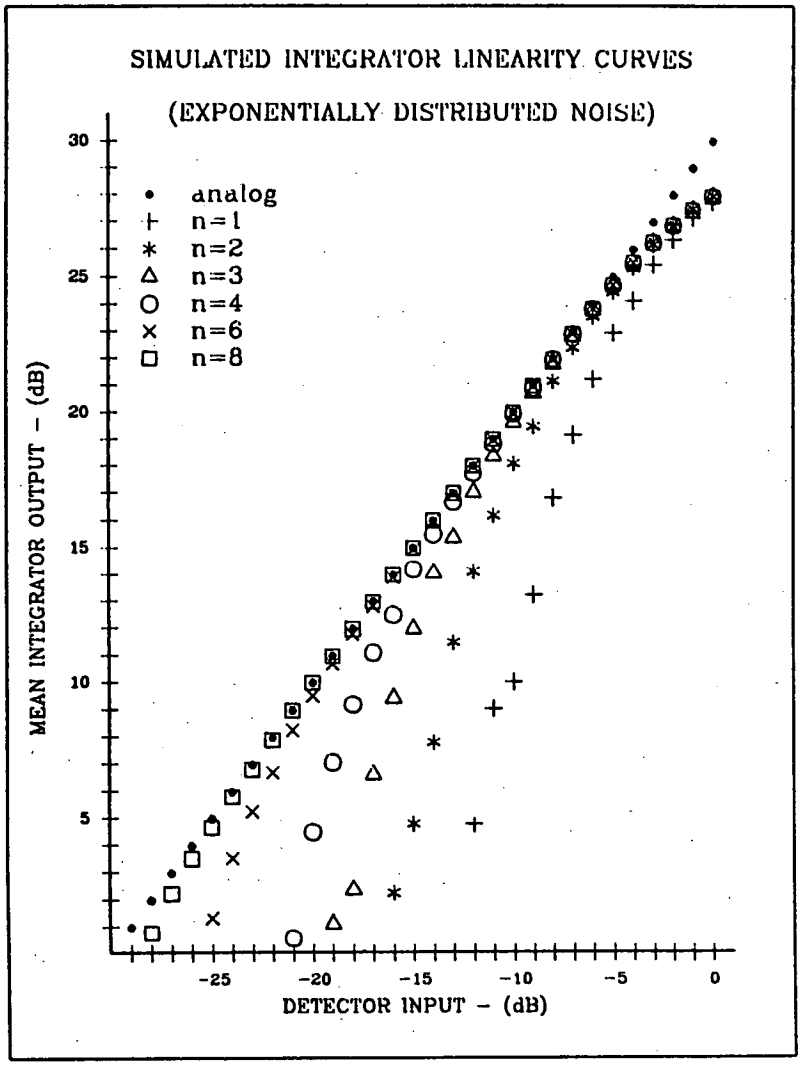
The linearity simulation involved two basic steps:

(a) generation of random numbers having either an exponential or a Rayleigh p.d.f. (see appendix three for details of this process);

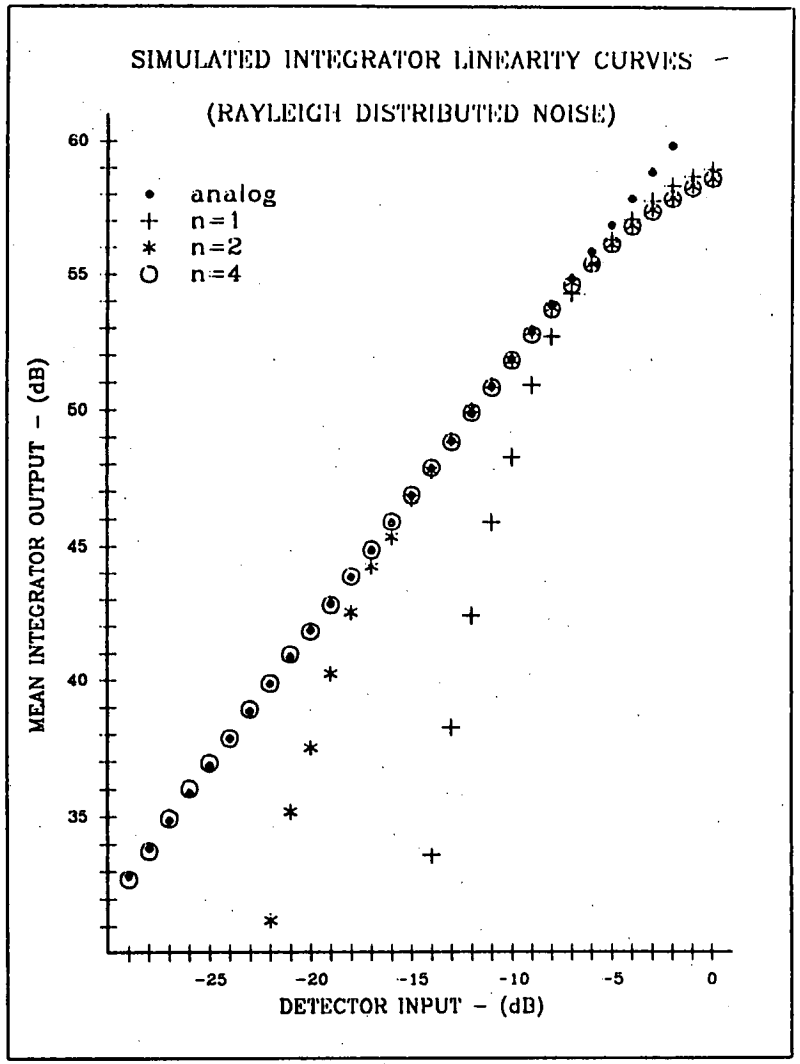
(b) addition of the numbers in the case of the analog system or quantization followed by addition in the other cases.

The same set of 1000 numbers was used for each input level and the transformation scale factors (equations III.2 and III.3) adjusted to provide the required mean input power (σ^2) to the detector. All systems (digital and analog) were presented with the same overall set of random numbers, eliminating statistical fluctuation resulting from the use of different data sets. However, a check on the magnitude of fluctuations due to the use of different data sets indicated that changes in the computed characteristics were negligible (~0.1 dB).

The results are shown in figures 5-3(a) and 5-3(b) for analog and n-bit digital systems. The horizontal axis in both graphs is



(a)



(b)

Fig.5-3: Linearity curves for simulated analog and n-bit digital systems. Curves for a receiver using a full-wave square-law detector are shown in (a); full-wave linear (modulus) detector curves are shown in (b).

labelled in terms of mean power at the detector input with 0 dB corresponding to $\sigma^2 = 1$, the value giving a mean quantizer input of 1.0 V with a square-law detector. Detector transfer functions of the form $y = b^2$ and $y = |b|$ are assumed. Spectral power is the required quantity regardless of detector type, so the sum derived from 1000 Rayleigh variates was squared prior to plotting, causing the ordinates for the linear detector graph to be larger than those for the square-law plot. Expressed in decibels, the integrator output for each system at each input level is

$$x_{(dB)} = 10 \log x \quad , \qquad 5.3.1$$

where 'log' means base ten logarithm and

$$x = \sum_{i=1}^{N} a_i \qquad 5.3.2$$

(a_i are the random numbers in the analog system or their quantized representations in the digital systems).

At high input levels the signal is clipped severely by the quantizer, causing the integrated output to rise asymptotically to a limit of N × 1.0 V, corresponding to 30 dB and 60 dB for the square-law and linear cases respectively. As the input is reduced the quantizer begins to operate in the optimum region and the digital systems closely approximate the analog. Further reduction causes the mean of the detector output p.d.f. to fall well below the first decision level, so the quantizer ceases to respond to many excursions of the input signal. In this region digital systems again show non-linearity.

The observed variation in x with different overall random number sets is not surprising. For large N, conventional sampling theory and error propagation analysis (see, for example, Spiegel, 1975) can be used to show that the fractional error in x is

$$\varepsilon_1 \sim 1/\sqrt{N}$$

for both linear and square-law detectors. Denoting the error in

simulation by $\Delta x_{(dB)}$,

$$x_{(dB)} \pm \Delta x_{(dB)} \sim 10 \log(x) \pm 10 \log (1+\varepsilon_1) \qquad 5.3.3$$

if $\varepsilon_1 << 1$. For $N = 1000$, $\Delta x_{(dB)} \sim 0.1$, the observed value.

In simulating the SNR performance of the various systems the SNR was defined as:

$$SNR = \bar{x}/s, \qquad 5.3.4$$

where $\bar{x}$ is the mean integrated output for each system at each level; that is, the mean of many values of x produced by using different random number sets. The standard deviation of the many values of x is denoted by s. For the linear detector $\bar{x}$ and s were computed after the post-integration squarer.

After M estimates of x,

$$\bar{x} = (1/M) \sum_{j=1}^{M} x_j \qquad 5.3.5$$

and

$$s = \sqrt{\frac{\sum_{j=1}^{M} (x_j - \bar{x})^2}{M-1}} \; . \qquad 5.3.6$$

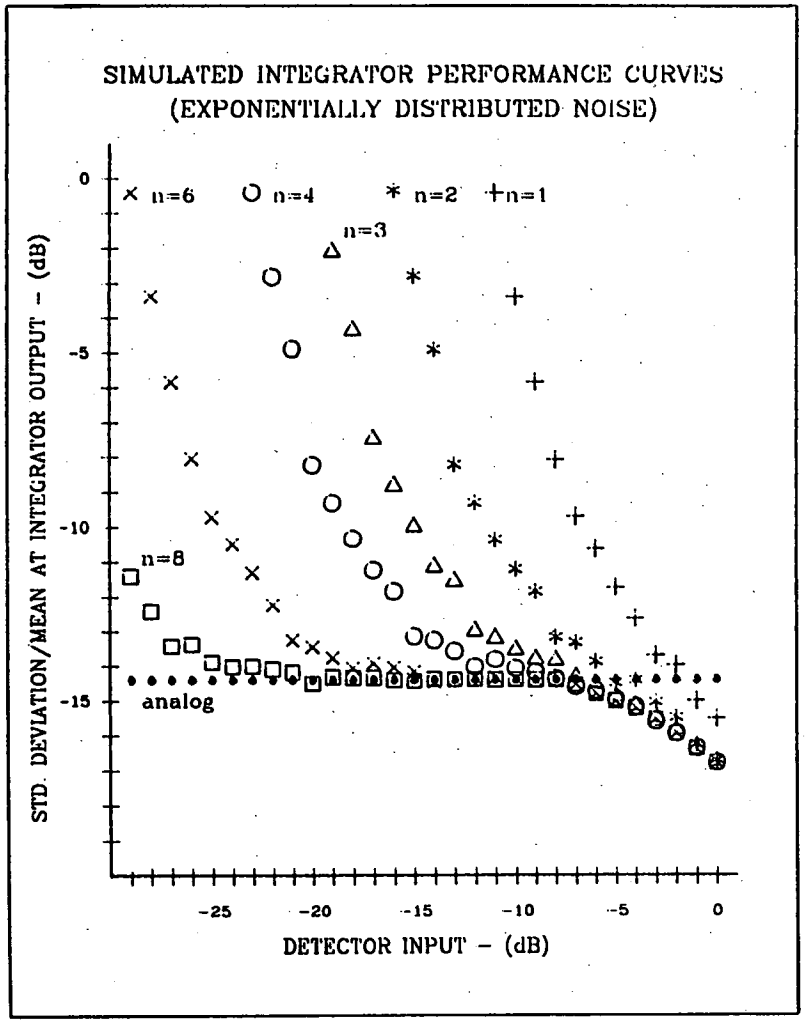
Expressed in decibels,

$$SNR_{(dB)} = 10 \log (SNR).$$

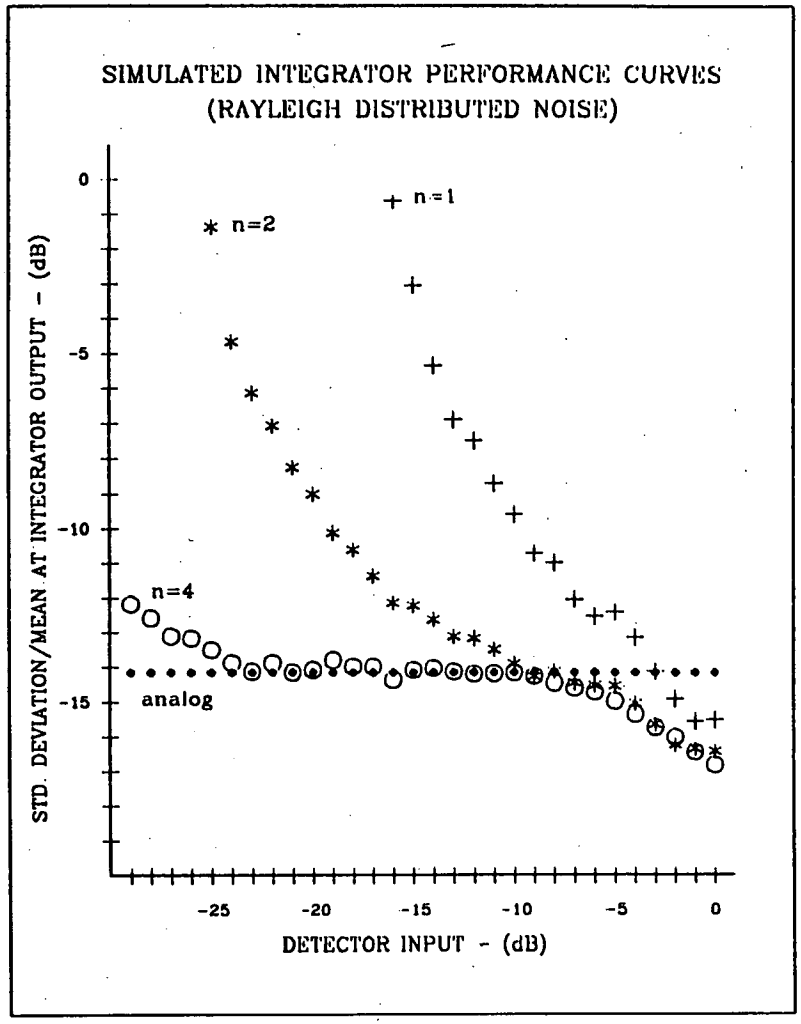
In radioastronomy the frequently used form of the radiometer equation yields the noise-to-signal ratio (NSR) of the integrated output. In the present study this quantity has been plotted directly. Of course,

$$NSR_{(dB)} = - (SNR)_{(dB)} \; .$$

The results of the simulation are plotted in figures 5-4(a) and 5-4(b). At high input levels severe clipping causes the difference $x_j - \bar{x}$ in equation 5.3.6 to be small, so as the level is increased the



(a)



(b)

Fig.5-4: Noise-to-signal ratio curves for simulated analog and n-bit digital systems. The curves in (a) are for a receiver using a full-wave square-law detector while the curves in (b) apply to a receiver having a modulus detector.

quantity $NSR_{(dB)}$ becomes more negative. In practice, non-linearity prevents operation in this range. At very low input levels the quantizer fails to respond to most of the signal variations so the SNR is degraded severely by quantizing effects.

Each point on the graphs was obtained with N = 1000 and M = 50. A different set of 1000 numbers was used for each of the 50 data sets, so each point is the result of manipulating 50 000 random numbers. The same 50 000 numbers were used (with different scale factors) for each input level and all systems simulated were presented with the same overall data set to eliminate fluctuations due to different data.

As in the linearity investigation, conventional statistical methods can be used to estimate simulation errors. Assuming x_j to be normally distributed the fractional error in $R = s/\bar{x}$ can be shown to be

$$\varepsilon_2 \sim \sqrt{\frac{R^2 + \frac{1}{2}}{M}} \quad .$$

From equation 5.3.3,

$$R_{(dB)} \pm \Delta R_{(dB)} \sim 10 \log (R) \pm 10 \log (1 + \varepsilon_2).$$

For N = 1000 and M = 50, $\varepsilon_2 \sim 0.1$ and $\Delta R_{(dB)} \sim 0.5$ for both types of detector.

Within the ± 0.5 dB error in estimating the position of the analog reference line, the curves reflect accurately the relative performance of continuous and quantized systems. To check the validity of the error estimation, ten separate evaluations of the NSR for the analog integrator were made using entirely different sets of random data. For N = 1000, M = 50 the results were:

$$NSR_{sq.law} = -15.00 \pm 0.44 \quad dB$$

$$NSR_{linear} = -14.70 \pm 0.46 \quad dB.$$

As the expected NSR with a square-law detector is $1/\sqrt{1000}$ or -15.00 dB, the simulation gives plausible results in an absolute sense and the analog reference lines in figures 5-4(a) and 5-4(b) show, coincidentally, worst-case deviations from the average positions.

The final simulation was an extension of the check just described. The number of samples (M) used to estimate the NSR for the analog integrator was increased to 10 000 and ten separate NSR evaluations made using different data sets for $1 < N < 2048$. At each value of N, the average SNR difference between systems with square-law and linear detectors was calculated and the spread in the ten calculated differences retained. Figure 5-5 shows the SNR advantage factor for the square-law detector plotted as a function of $\log_2 N$. The error bars indicate 95% confidence intervals derived using small-sample theory (t-statistics) and by treating the two SNR values from each of the ten simulations as paired data. Error bars are not shown for the $N = 2^{12}$ (4096) point as only one simulation was performed for this value.

The results indicate that the square-law device always yields better SNR but that the difference between detectors never exceeds 0.2 dB. It is interesting to compare these results with those of Robertson (1968) (cited in Section 5.2) who examined the relative efficiency of linear and quadratic envelope detectors in detecting a CW signal corrupted by Gaussian noise in a communications circuit. A linear detector was superior for $N \lesssim 2^7$ post-detector samples but a quadratic detector yielded better results for large N. The maximum advantage reported in both regimes was about 0.2 dB, the square-law advantage reaching a large-N asymptote of slightly less than this value. Use of Robertson's technique may give a plot similar to figure 5-5 if his method were extended to cover low SNR cases where a very high probability of false alarm in the communications link is acceptable.

5.4 COMMENTS ON DIGITAL INTEGRATOR PERFORMANCE

Two main conclusions can be drawn from the simulations. The first concerns the quantizer resolution; the second the type of detector. Although dynamic range is not normally an important consideration in radioastronomy, a range of less than 6 dB or so is inconvenient in practice. In addition, if the spectrum from a SAW analyser is digitized between the -3 dB points, the output level in the resolution cells having the smallest relative sensitivity (those at the passband edges) must still be within the optimum input range of the quantizer. For a spectrometer using a square-law detector it is likely that the minimum

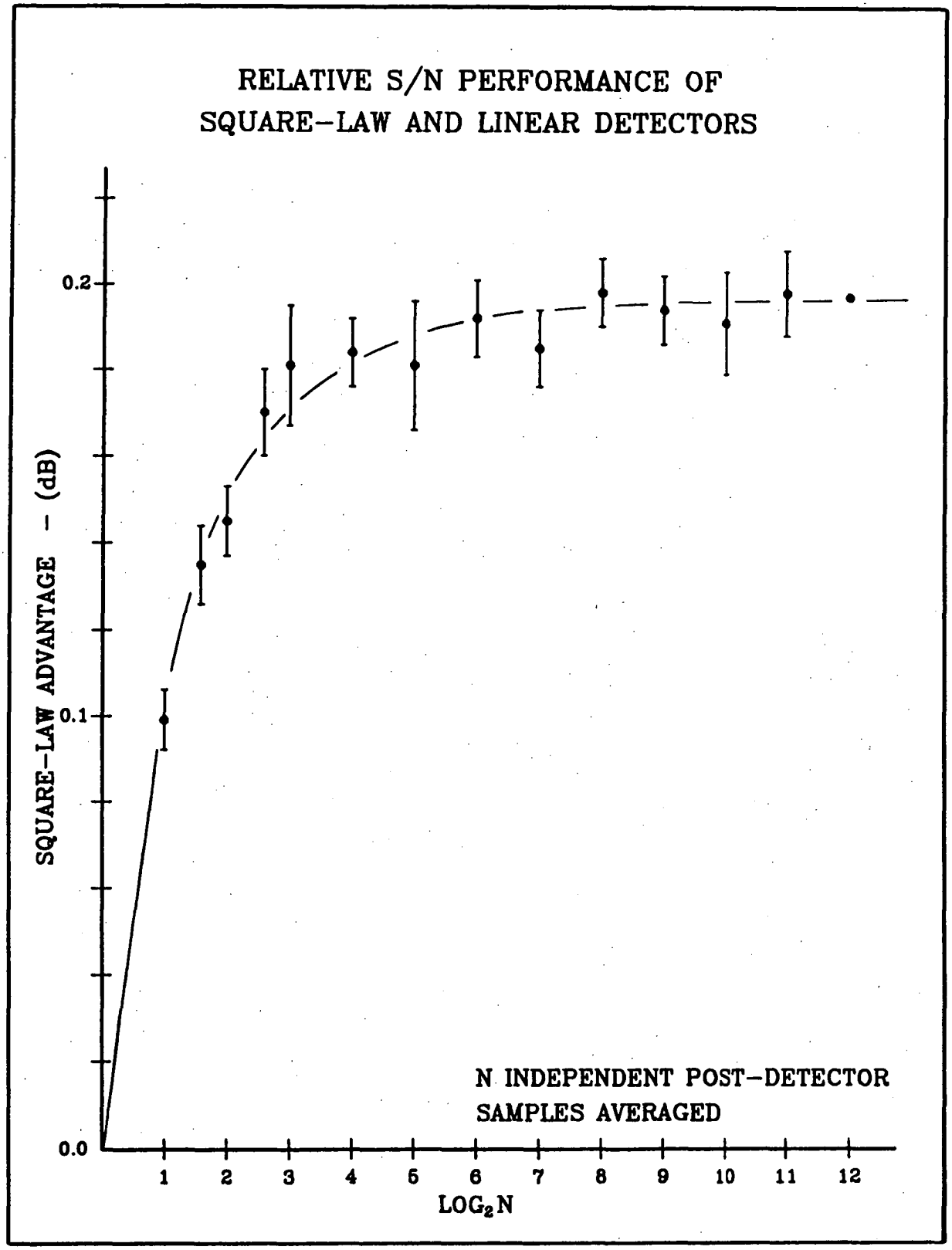


Fig.5-5: Relative signal-to-noise ratio performance of narrowband receivers having full-wave square-law and linear detectors. The square-law detector advantage is plotted as a function of $\log_2 N$. Error bars show 95% confidence limits. No error bars are shown for the $N = 2^{12}$ point since the result is from 1 simulation only.

practical quantizer resolution is four bits, giving a dynamic range of about 6 dB. Higher resolution quantizers may be justified in some circumstances but, in practice, the choice of video A/D converters is not arbitrary since cost escalates rapidly with increasing resolution while conversion speed decreases. With current technology it is difficult to make A/D converters capable of quantizing to more than eight-bit resolution and converting faster than 30 mega-samples/second.

Secondly, the investigations show that a square-law detector is not essential in all applications. A linear detector gives significantly more dynamic range with only a slight SNR penalty (0.2 dB loss corresponds to 95% radiometer operating efficiency). There are many occasions in astronomy when the SNR is of prime concern but it is worth remembering that a small variation away from the optimum square-law detector level is sufficient to lose 0.2 dB SNR in a low resolution post-detection quantizer. With computer controlled calibration methods the traditional objection to linear detectors (the tedium of establishing an absolute scale) is removed and, in any case, for a small change in antenna temperature (ΔT) around a large background level (T) the detector is approximately linear in ΔT since $\sqrt{T + \Delta T} \sim \sqrt{T}\,(1 + \frac{1}{2}\Delta T/T)$.

5.5 DIGITAL INTEGRATION OF ANALOG CORRELATOR OUTPUTS

The analogy between one SAW spectrum analyser resolution cell and a narrowband channel can be extended to cover the case where the outputs of two SAW analysers are processed using a correlator. The analogous system is shown in figure 5-6. Both correlator input signals are assumed to be Gaussian disturbances with zero mean. The post-correlation filter passes the video spectrum intact but rejects RF components. In the special case where the inputs are completely correlated the problem of determining the p.d.f. at the filter output simplifies to the square-law detector case already discussed.

When the multiplier inputs are not completely correlated, the situation is difficult to treat analytically. In a classic paper, Lampard (1956) analysed the case of an analog correlator having no video filter and gave very basic expressions for the p.d.f. when a video filter is used. With no filter the p.d.f. at the multiplier output is

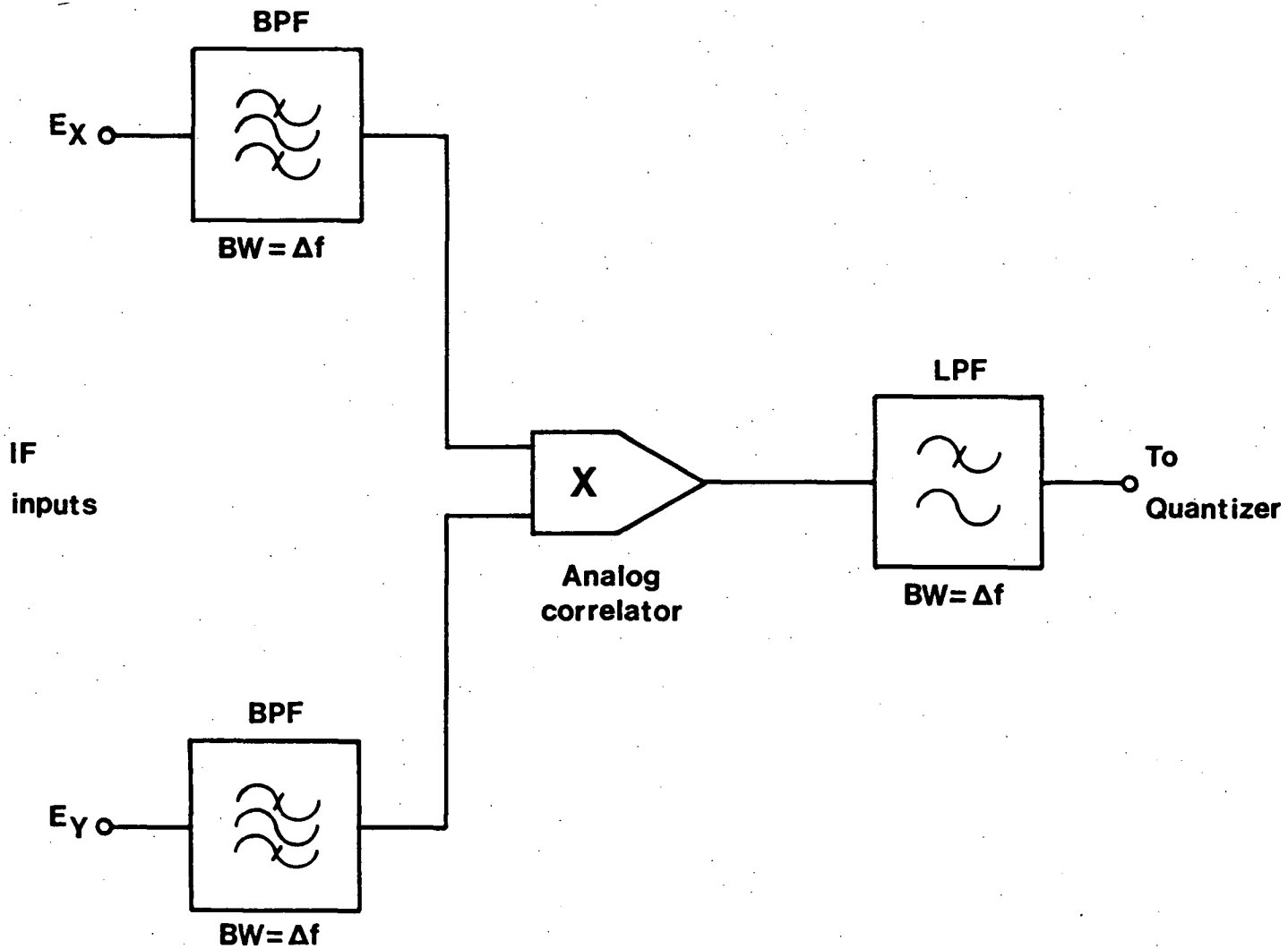


Fig.5-6: Narrowband correlation system.

$$p(w) = \frac{1}{\pi\sqrt{1-\rho}} \exp\left(\frac{\rho w}{1-\rho^2}\right) K_0\left(\frac{w}{1-\rho^2}\right) , \qquad 5.5.1$$

where ρ is the correlation coefficient at the multiplier inputs and $K_0(x)$ is the modified Bessel function of the first kind of order zero. Cole (1968) shows that the mean of the p.d.f. is

$$\bar{w} = \rho \qquad 5.5.2$$

and the NSR at the correlator output is

$$\frac{\sigma_{out}}{\rho} = \frac{\sqrt{1+\rho^2}}{\rho} . \qquad 5.5.3$$

The impediment in simulating a digital post-correlation integrator is the lack of an expression analogous to equation 5.5.1 when a video filter is used.

A few qualitative observations about equation 5.5.1 can be made. For $\rho = 0$, $p(w)$ is just $K_0(w)/\pi$. At $\rho = 1$, $p(w)$ reduces to the un-filtered square-law detector density function. This function is a χ^2 p.d.f. having one degree of freedom and tends to infinity for small w. It would be difficult therefore to quantize the un-filtered output effectively. When RF terms are removed, the resulting p.d.f. has two degrees of freedom (section 5.2). As the video bandwidth is decreased, more detector noise pulses are averaged, the number of degrees of freedom increases and the output p.d.f. becomes more like a Gaussian function. A similar situation is expected with the general ($\rho \neq 1$) case. It is likely that the presence of any video filtering makes the problem of quantizing the correlator output easier from a practical signal processing viewpoint. As the video bandwidth is decreased quantization becomes straightforward because, by the central limit theorem, the filter output p.d.f. is asymptotically normal.

CHAPTER 5 - REFERENCES

Cole, T.: Finite sample correlations of quantized Gaussians. Aust. J. Phys., Vol.21, 1968, pp.273-282.

Cooper, B.F.C.: Autocorrelation spectrometers - in "Methods of experimental physics" (M.L. Meeks ed.). Vol.12, Ch.3.5, Academic Press, 1976.

Davenport, W.B. and Root, W.L.: "An introduction to the theory of random signals and noise". Ch.12, McGraw-Hill, 1958.

Emerson, R.C.: First probability densities for receivers with square-law detectors. J. App. Phys., Vol.24, No.9, Sept. 1953, pp.1168-1176.

Kac, M. and Siegert, A.J.F.: On the theory of noise in radio receivers with square-law detectors. J. App. Phys., Vol.18, No.4, April 1947, pp.383-397.

Lampard, D.G.: The probability distribution for the filtered output of a multiplier whose inputs are correlated, stationary, Gaussian time series. IRE Trans. Inf. Theory, Vol. IT2, No.1, 1956, pp.4-11.

Meyer, M.A. and Middleton, D.: On the distribution of signals and noise after rectification and filtering. J. App. Phys., Vol.25, No.8, August 1954, pp.1037-1052.

Middleton, D.: "An introduction to statistical communication theory". Ch.2.2, McGraw-Hill, 1960.

Robertson, G.H.: Performance degradation by post-detector non-linearities. Bell Syst. Tech. J., Vol.47, No.3, Mar. 1968, pp.407-414.

Spiegel, MR.: "Theory and problems of probability and statistics". Ch.5, McGraw-Hill, 1975.

Whalen, A.D.: "Detection of signals in noise", Ch.4, Academic Press, 1971.

CHAPTER 6 SAW SPECTROMETER DESIGN

6.1 INTRODUCTION

In this chapter some practical aspects of SAW spectrometer design are discussed. The section dealing with coherent analyser implementation is original, while little of the discussion relating to fundamental design techniques has been published elsewhere. Although it is assumed that the spectrometer is based on the (C)-M-C or compressive receiver configuration, the discussion extends readily to other SAW processor architecture.

The chapter begins with a discussion of DDL selection criteria. Basic RF design procedures for compressive receivers are then examined. In the remaining sections of the chapter the practical realization of a coherent IF processor (figure 4-10) is considered. For descriptive purposes the instrument is divided into four parts. These are:

(a) the impulse generator;
(b) the pre-multiplier (LO) chirp generation chain;
(c) the input signal processing, signal mixing and convolution stages;
(d) the post-transformation signal splitting and processing stages.

Much of the discussion is equally relevant to single-channel SAW analysers.

6.2 CHOICE OF SAW DISPERSIVE DELAY LINES

In principle, DDL parameters are determined by the full-compression bandwidth (B_1) and frequency resolution (Δf) required in the SAW compressive receiver. These two specifications give the convolver time-bandwidth product in an un-weighted instrument since $T_1B_1 = B_1/\Delta f$. Note that the chirp time (T_1) of the convolver is the reciprocal of the frequency resolution. If a weighted convolver is used, T_1 needs to be extended (with a corresponding increase in T_1B_1) by a factor equal to the compressed pulse broadening factor if frequency resolution equal to that of an un-weighted analyser is to be retained.

The maximum usable convolver TB is likely to be limited by sampling speed and data buffering capabilities of digital post-transformation processors. In applications outside astronomy the SAW spectrometer analysis time of either T_1 or $2T_1$ (depending on the duty cycle) may also be a limitation, again restricting the maximum value of T_1B_1. Assuming T_1B_1 to be specified, the required LO time-bandwidth product is $4T_1B_1$. The largest in-line SAW DDL's give TB~1000 so if different, in-line devices are used for expansion and compression a practical limit for T_1B_1 is 250. When RAC devices are used the maximum values of TB are increased by an order of magnitude.

If the required T_1B_1 is within the available TB range but $4T_1B_1$ is not, the LO chirp can be generated using a combination of two devices each with $TB = T_1B_1$. In fact, identical DDL's can be used in an expansion and compression role if spectral inversion (chirp slope negation) is feasible at some point in the analyser (section 6.5).

DDL's obtained for use as convolvers should not be reciprocal ripple compensated since in a compressive receiver the Fresnel characteristic of the long-duration chirp entering the convolver is not matched to the convolver ripple compensation. Furthermore, the Fresnel ripple pattern on the portion of the signal chirp entering the convolver varies when the analyser input frequency is altered, so the problem of excess sidelobes (section 4.4) which exists in any SAW analyser is accentuated greatly if a compensated convolver is used.

The choice of DDL substrate material and device encapsulation is usually made by the manufacturer, but purchasers of SAW devices for spectrum analyser applications must be discerning in choosing between potentially suitable devices. For example, the temperature sensitivity of the analyser (section 4.5) should be estimated and a decision made concerning the necessity or feasibility of stabilizing the DDL thermal environment. Choosing a production device with a less temperature-sensitive substrate will minimise stabilization problems although the deleterious effects of the inescapable increase in DDL insertion loss must be weighed against the advantage.

Additionally, some radar DDL's exhibit a fairly high degree of electrical coupling between the input and output ports. In the intended application the device outputs are time gated to eliminate feed-through of excitation transients but in SAW spectrum analysers, particularly instruments operating at high duty cycle, gating may be impossible and the valid output data window may be corrupted by low-level feed-through. If possible, select DDL's with high isolation together with absolute delay times which ensure that spurious transients appear outside the data window at the operating duty cycle of the analyser.

6.3 BASIC COMPRESSIVE RECEIVER RF DESIGN

When a pulse expansion DDL having an impulse response of duration T is excited with a pseudo-impulse (a practical approximation to a delta function) of duration $\tau = 1/B$, conservation of signal energy at the device input and output results in the equality:

$$V_i^2 \tau = V_o^2 T$$

or

$$V_o = \frac{V_i}{\sqrt{TB}} \quad , \qquad 6.3.1$$

where V_i and V_o are the input and output signal amplitudes, and the impedance at the input and output is assumed the same. Bandwidth restrictions inherent in practical DDL's ensure that the effective pseudo-impulse duration is about 1/B, even with no external band-limiting.

From the previous expression, the input and output peak power levels are related by the simple scaling,

$$P_o = \frac{P_i}{TB}$$

or

$$P_o(dB) = 10[\log P_i - \log(TB)].$$

An expansion loss (EL) is defined such that

$$EL = 10 \log(TB) \quad dB. \qquad 6.3.2$$

This is the amount by which the peak power is reduced at the dispersive filter output if device insertion losses are ignored. Similar reasoning for a matched compression filter leads to an identical definition of compression gain (CG). That is,

$$CG = 10 \log(TB) \quad dB. \qquad 6.3.3$$

If the convolver DDL is weighted to reduce sidelobes, it is no longer matched exactly to the applied chirp signals and the compression gain is reduced slightly. The weighting lengthens the duration of compressed pulses, leading to a degradation in spectrum analyser frequency resolution (section 3.5). For a weighted convolver the compression gain is

$$CG_W = [10 \log (TB)] - M \quad dB, \qquad 6.3.4$$

where M is the mismatch loss. For most weighting functions M < 1 dB.

Figure 6-1 shows a basic compressive receiver in which the required factor of four in TB between the LO and convolution chirps is obtained by using DDL's with appropriate specifications. In some spectrum analysers more elaborate techniques are used to generate the LO chirp and these alternative methods are considered in section 6.5. The analysis of the simple receiver can be extended easily to more complex schemes, so a detailed examination of the basic configuration is worthwhile. The method of analysis is similar to that used by Davie and Arthur (1979).

In figure 6-1 B_1 = 30 MHz, T_1 = 2.4 μs and T_1B_1 = 72. These parameters are based on those of SAW devices used in the prototype compressive receiver described in chapter seven. Both signal and noise paths are shown in figure 6-1 and representative values of noise factor (F), conversion loss (CL) and insertion loss (IL) are assumed. Noise is assumed to be thermal in character with an initial power given by

$$P_N = k\theta B \qquad \text{watt}$$

or

$$P_N = [10 \log B_{(MHz)}] - 114 \quad dBm,$$

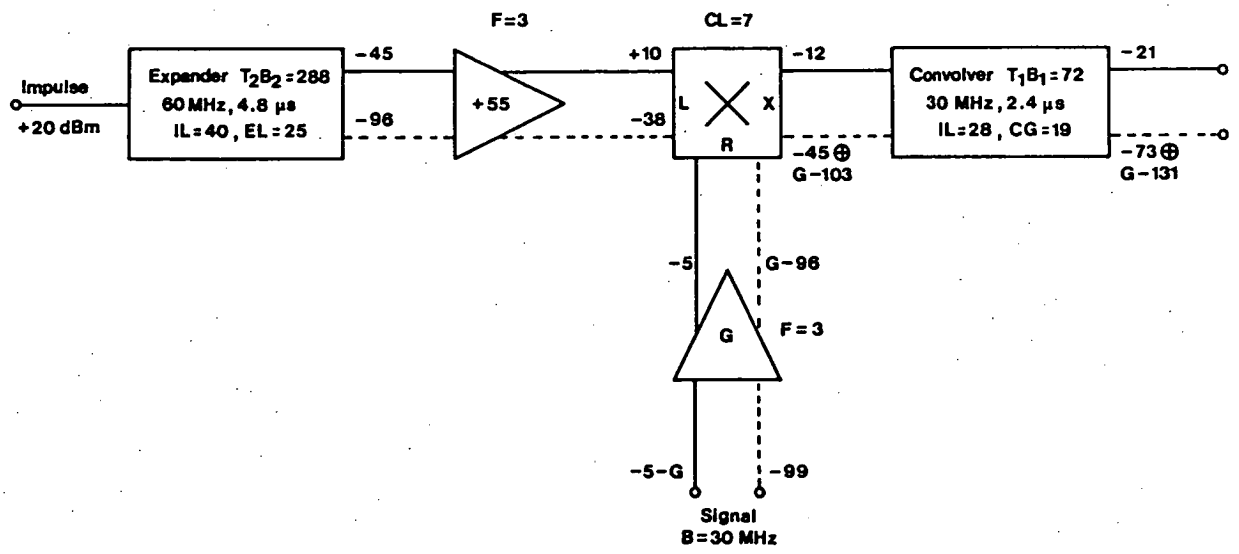


Fig.6-1: Basic compressive receiver analysed in the RF design example. Signal paths are shown as solid lines; noise paths are dotted. Signal path levels are maximum values (in dBm). A level marked X⊕Y is either X or Y, whichever is the higher. When X = Y the level is X (or Y) + 3 dBm.

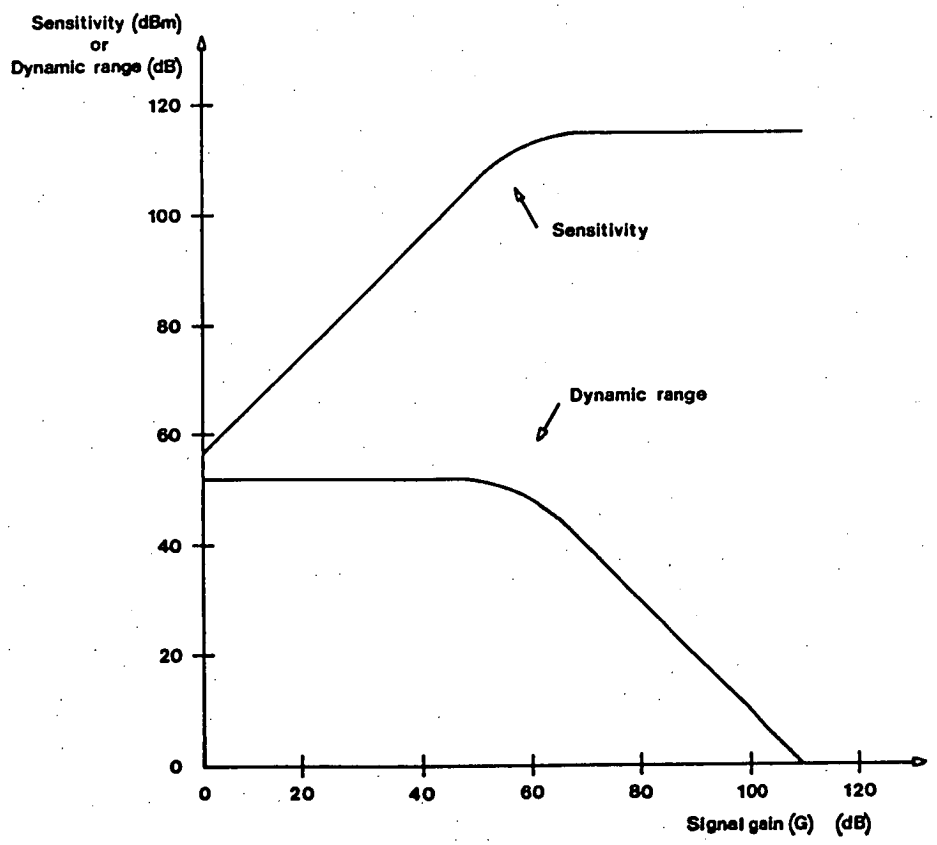


Fig.6-2: Sensitivity and dynamic range of the compressive receiver of figure 6-1 plotted as a function of G, the signal gain. The sensitivity is read by multiplying the ordinates by -1.

where k is Boltzmann's constant and θ, the absolute temperature, is taken as 290 K in the second expression.

The noise transfer through the double balanced mixer is calculated by assuming that a noise disturbance superimposed on the chirp at the LO (L) port produces a perturbation at the IF (X) port of similar magnitude to that which would occur if the same disturbance were injected at the signal (R) port. The implication of this noise transfer model is that the noise floor at the mixer output is effectively determined by either LO noise or noise from the signal amplifier, whichever is greater. When the noise power at the L and R ports is equal, the X port signal-to-noise ratio is 3 dB worse than the SNR at the R port.

It is assumed that image band noise entering the convolver is negligible. This is achieved in practice by restricting the signal amplifier bandwidth or by using an SSB mixer. If the amplifier image noise is not restricted, the analyser output noise floor when determined by the signal chain is raised 3 dB.

The LO amplifier gain must be sufficient to ensure that the mixer L port is driven to saturation, thereby ensuring that conversion loss and distortion specifications are met. The quoted drive level of +10 dBm in figure 6-1 is a typical value for diode mixers. The signal path gain is left as a variable to illustrate the eventual compromise necessary between spectrum analyser sensitivity and dynamic range. The only constraint on the gain is that the maximum signal level at the mixer R port should not exceed about -5 dBm for good mixer operation. In practical instruments post-convolver amplifiers are necessary but inclusion of such stages does not alter materially the spectrum analyser characteristics.

The dynamic range of the analyser is defined as the difference between the maximum output level and the noise floor. The instrument sensitivity is defined as the input level needed to produce an output level equal to the noise floor. The best dynamic range of the spectrum analyser in figure 6-1 is limited to 52 dB by virtue of the -96 dBm of noise power at the expansion DDL. If the signal gain (G) is small this is the actual dynamic range but, as G is increased, noise from the signal

path becomes significant. Dynamic range and sensitivity are both plotted as a function of signal gain in figure 6-2.

At G = 58 dB noise contributions from the L and R ports are equal, giving an output noise power of -70 dBm, a dynamic range of 49 dB and a sensitivity of -112 dBm. At G = 110 dB, R port noise is entirely dominant, leading to an output noise floor of -21 dBm. At this point the dynamic range is zero with a sensitivity limit of -115 dBm. The optimum value of G is that which produces approximately equal noise power contributions at the L and R ports and, for the spectrum analyser illustrated, G = 58 dB is the best choice. With a knowledge of typical signal level requirements in a SAW analyser, the individual modules can be considered in detail.

6.4 IMPULSE GENERATION

The impulse generator supplies the short pulse, or pseudo-impulse, necessary to excite the expansion DDL. A pulse having either polarity is acceptable. Ideally, the pulse spectral density should be constant across the DDL passband although it is permissible to rely on the limiting action of the mixer to level the LO amplitude if the variation across the chirp waveform is not more than a few decibels.

The analysis of the previous section showed that LO SNR is a limiting factor in SAW spectrum analyser performance. It is therefore important that enough impulse power be available to ensure that the best possible SNR is obtained. Typical DDL specifications allow peak excitation signals of between one and ten volts, corresponding to a peak impulse power range of +12 dBm to +32 dBm with a 50 ohm load.

Some means of repetitively triggering the impulse generator must be provided and, in a compressive receiver designed to be interfaced to digital systems, it is convenient to make the triggering circuitry compatible with the logic family used in the timing control circuitry. The impulse generator must trigger with a minimum of jitter and must be well shielded and de-coupled to prevent coupling of the impulse into low-level sections of the spectrum analyser and post-transformation signal processors.

There are two basic approaches to impulse generator design. The first is to generate the high-level pulse directly; the second is to use a low-level pulse generator followed by several wideband amplifiers. If the second method is used, a small RF power amplifier needs to be included as an output stage. It is difficult to amplify faithfully very short pulses because of the transient performance limitations of many RF amplifiers. In particular, the time domain characteristics of feedback amplifiers often render them unsuitable for a pulse amplification role. Since direct generation of the impulse is practical and is simpler than the alternative scheme, direct generation is preferred.

Two types of direct generation impulse generator have been constructed and used successfully in prototype SAW spectrum analysers. The first type is based on a VMOS power MOSFET switching a low-voltage source to the input of a shorted transmission line. The differentiator action of the shorted line (Coekin, 1975) produces a pulse of width equal to the total propagation time on the line. In a prototype generator, 0.6 m of miniature coaxial cable (velocity factor ~ 0.67) was used as a delay line, leading to an expected pulse width of 6 ns. The actual width at half height was around 10 ns, the excess being attributed to switching speed limitations in the MOSFET.

The second impulse generator is derived from a design described by Kowatsch and Lafferl (1980). The circuit, shown in its basic form in figure 6-3, relies on the avalanche multiplication effect in the collector-base space charge layer of epitaxial planar silicon transistors. The supply voltage V_s and base resistor R_b are chosen to give a stable operating point with a collector-emitter voltage (V_{CE}) near the collector-base breakdown voltage (V_{CBO}). The quiescent collector current is very low and in this condition capacitor C_1 is charged to very nearly V_s.

When a positive trigger pulse (TTL level compatible) is applied to the base the transistor turns on, causing C_1 to discharge with a high current impulse through R_1 ($R_1 << R_c$). Due to the extreme rapidity of the avalanche effect, the switching time is of the order of one nano-second for most small signal transistors. For a given load the output pulse width and height depends on C_1. With the component values

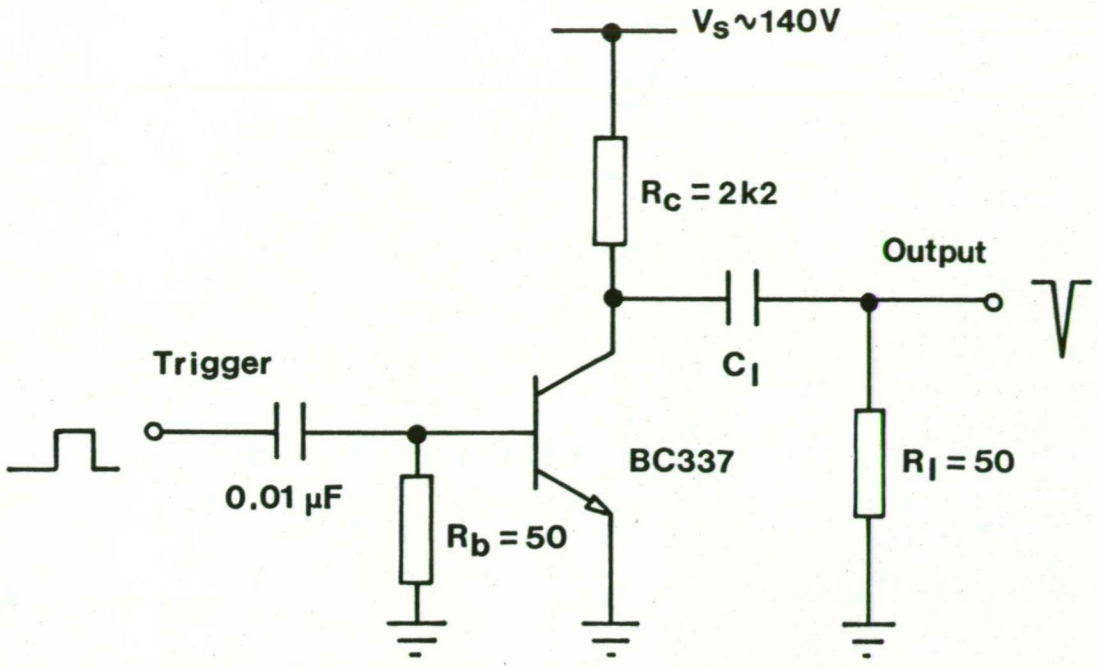


Fig.6-3: Avalanche effect impulse generator.

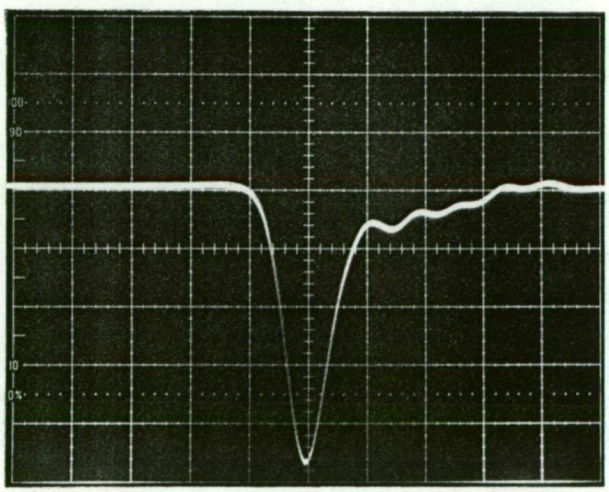


Fig.6-4: Typical output from the circuit of figure 6-3 (C_ℓ=47pF) after attenuation to suit SAW dispersive delay lines. Horizontal scale: 2 ns/div. Vertical scale: 2 V/div.

shown, the output varies from 1.3 ns (measured at half-height) at 14 V to 3.8 ns at 70 V with C_1 = 5.6 pF and 82 pF respectively. Figure 6-4 shows a typical impulse generator output after attenuation to suit the requirements of SAW DDL's.

The generator triggering frequency is limited by the transistor power dissipation, the storage time (τ_{ST}) of minority carriers in the base region (hundreds of nano-seconds for typical transistors) and the charging circuit time constant, τ_{RC}. For $R_1 << R_c$ and low parasitic capacitance, $\tau_{RC} \sim R_c C_1$. About three time constants must elapse if C_1 is to recharge fully after triggering, so the minimum interval between trigger pulses is

$$T_{min} \sim \tau_{ST} + 3\tau_{RC} .$$

With typical components the maximum trigger frequency is about 1 MHz, sufficient for most SAW spectrum analyser applications. For higher trigger frequencies the simplicity of the basic generator makes it feasible to construct parallel units triggered independently from a poly-phase trigger source. The outputs of the separate units can then be combined via isolating diodes to yield the required excitation pulse train.

In addition to being simple, the avalanche impulse generator has the advantage of producing high-level pulses. The attenuation network needed to reduce the pulse amplitude to one compatible with SAW devices defines a resistive output impedance which is independent of frequency or drive level. The resistive output minimises interaction between the generator and the expander DDL internal impedance matching networks. This type of interaction is a common cause of excess ripple on chirp waveforms.

Although a single pulse is often used to excite an expansion DDL, for a fixed pulse height it is possible to obtain more power around the device centre frequency by using an excitation waveform consisting of several pseudo-impulses separated in time by the reciprocal of the centre frequency. Real impulse generators have peak power or voltage limitations and DDL's have a maximum recommended excitation voltage.

Multiple impulse excitation affords a way of increasing the effective power without exceeding impulse generator or DDL specifications. The penalty is extra complexity in the impulse generator and an unavoidable variation in excitation power over the DDL passband.

The spectral density $S(\omega)$ of a single pulse is essentially flat over a bandwidth much narrower than the reciprocal of the pulse duration (T). More generally, the situation with an M-pulse train is analogous to an M-slit optical diffraction problem. The peak power density increases as M^2 or 20 log M dB and the shape of the spectral maxima varies according to the form $\sin^2(M\omega T)/\sin^2(\omega T)$ if the individual pulses are delta functions (Longhurst, 1973). As M increases, the width of the maxima decreases and the difference between the spectral density at the DDL centre frequency and that at the passband edges grows larger. In most SAW device applications, spectral narrowing makes it undesirable to use more than two or three pulses in the excitation waveform. Figure 6-5 shows the difference in power between the DDL centre frequency and band edges plotted as a function of device fractional bandwidth.

The avalanche effect impulse generator (figure 6-3) can be adapted readily to give multiple pulses if delayed versions of the trigger pulse are applied to separate generators. As mentioned, the outputs of individual generators can be combined using isolating diodes. The delay elements can be logic gates, coaxial or strip-line delay lines, or ECL multivibrators.

6.5 LO CHIRP GENERATOR DESIGN

The pre-multiplier (LO) chirp generation stages in a SAW compressive receiver provide a chirp signal having the same absolute slope as the convolver impulse response but having four times the time-bandwidth product. The sign of the LO chirp slope is such that when the LO is mixed with input signals, product chirps falling within the convolver passband have a slope which is the negative of the convolver impulse response slope (figure 4-7(a)).

There are a number of suitable positions in frequency for the LO chirp and the spectrum analyser input band. The input band can be

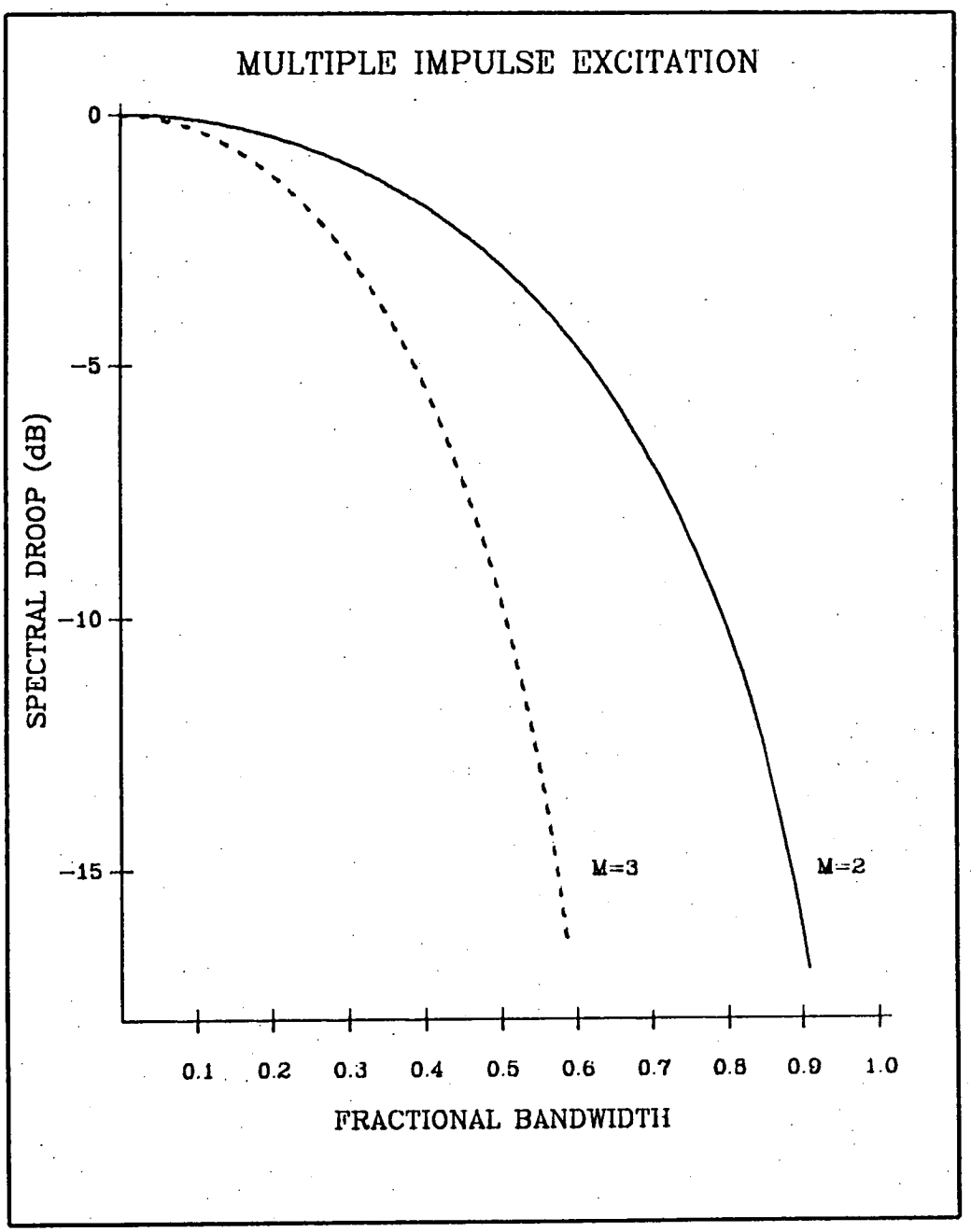


Fig.6-5: Spectral droop of a DDL excitation signal plotted as a function of the device fractional bandwidth for two and three-impulse waveforms. It is assumed that the first spectral maximum is placed at the DDL centre frequency.

translated up or down in frequency by mixing stages and the LO signal can be translated or spectrally inverted (chirp slope reversed) in similar stages. There is some advantage in having the input band, LO chirp and convolver passband separated in frequency, since any mixer break-through is then outside the convolver bandwidth. Although the convolver correlation process is highly effective in discriminating against signals other than the wanted product chirps, the passage of uncompressed signals through the convolver adds to the noise floor at the analyser output, thereby reducing the dynamic range.

Chirps within the convolver passband are produced at $f_{LO} \pm f_{signal}$ when a normal double balanced mixer is used, so it is possible for an image chirp to enter the convolver if the spectrum analyser input is not band-limited. However, the image does not correlate in the convolver since its chirp slope is the same as that of the convolver impulse response. In fact, the image suffers a 3 dB processing loss because it is dispersed to twice its initial duration by the DDL. Like spurious mixer signals, uncompressed image chirps add to the output noise floor, so it is desirable to band-limit the spectrometer input.

An example drawn from an actual spectrum analyser design (appendix five) will be useful in clarifying the LO frequency selection process. Identical down-chirp DDL's were chosen as expanders and as an un-weighted convolver. The bandwidth of the DDL's is 19 MHz and the impulse response duration is 34 μs, giving TB = 646. The centre frequency is 91.0 MHz. The requirement was therefore for an LO chirp with $T_2B_2 = 68 \times 38 = 2584$ to be mixed with a 19 MHz signal bandwidth to produce up-chirp signals in the range 91±9.5 MHz. In this case the centre frequency of the input band was set at 250 MHz by external constraints.

There were two possibilities for the slope and spectral position of the LO: either an up-chirp at (250+91)±19 MHz or a down-chirp in the range (250-91)±19 MHz. In the first case, lower frequency input signals produce compressed outputs emerging earlier in time; in the second case the input band appears inverted at the spectrum analyser output because higher frequency signals at the analyser input produce earlier compressed pulses. In the actual design, the type of frequency translation equipment available meant that the 159±19 MHz position was preferred. The resulting output inversion is not a problem since it can be negated by either inverting the input band in an earlier mixing stage or, if a digital

post-transformation processor is used, reading the data buffer contents in reverse order.

The frequency-time diagram for the analyser is shown in figure 6-6. Two fixed oscillators at 58.5 MHz and 77.5 MHz are used to translate the expander outputs to 159±19 MHz. The mechanism for joining the two chirps is discussed shortly. The image band is at (159-91)±9.5 MHz and is easily separated from the wanted 250±9.5 MHz (full-compression) band by a simple filter. Note that the input band, LO chirp and convolver passband are all separated.

When expansion DDL's having the same time-bandwidth product as the convolver are used (as in the example), a double duration, double bandwidth LO chirp must be formed from the component chirps. There are two basic methods for deriving the chirp and each will be described briefly.

The first chirp generation system is the one implied in the example and illustrated in figure 6-7. This arrangement is referred to as an active chirp generator. Two expander DDL's and associated impulse generators are used. The impulse generators are triggered alternately with a pulse repetition interval of $T_2 = 2T_1$. The DDL outputs are gated to restrict the output to the required time slot, precluding direct impulse feed-through. Each gated output is mixed with a synthesized translation oscillator and the two mixer outputs summed to produce a continuous chirp at the required centre frequency. The gating elements can be double balanced mixers or PIN diode switches and an in-phase hybrid can be used as a summing device.

The method relies on stable fixed oscillators and trigger signals. Satisfactory operation is obtained when a crystal oscillator is used as a reference for both the phase-locked translation oscillators and the trigger source. It is desirable to have the facility to move one trigger or gating signal with respect to the other to minimise any chirp discontinuity. This facility can be obtained by using logic gates as signal delay elements.

A passive method of producing an LO chirp is shown in figure 6-8. Two expansion DDL's are cascaded to produce a double duration chirp. This chirp is then frequency-doubled to give a chirp having a time-

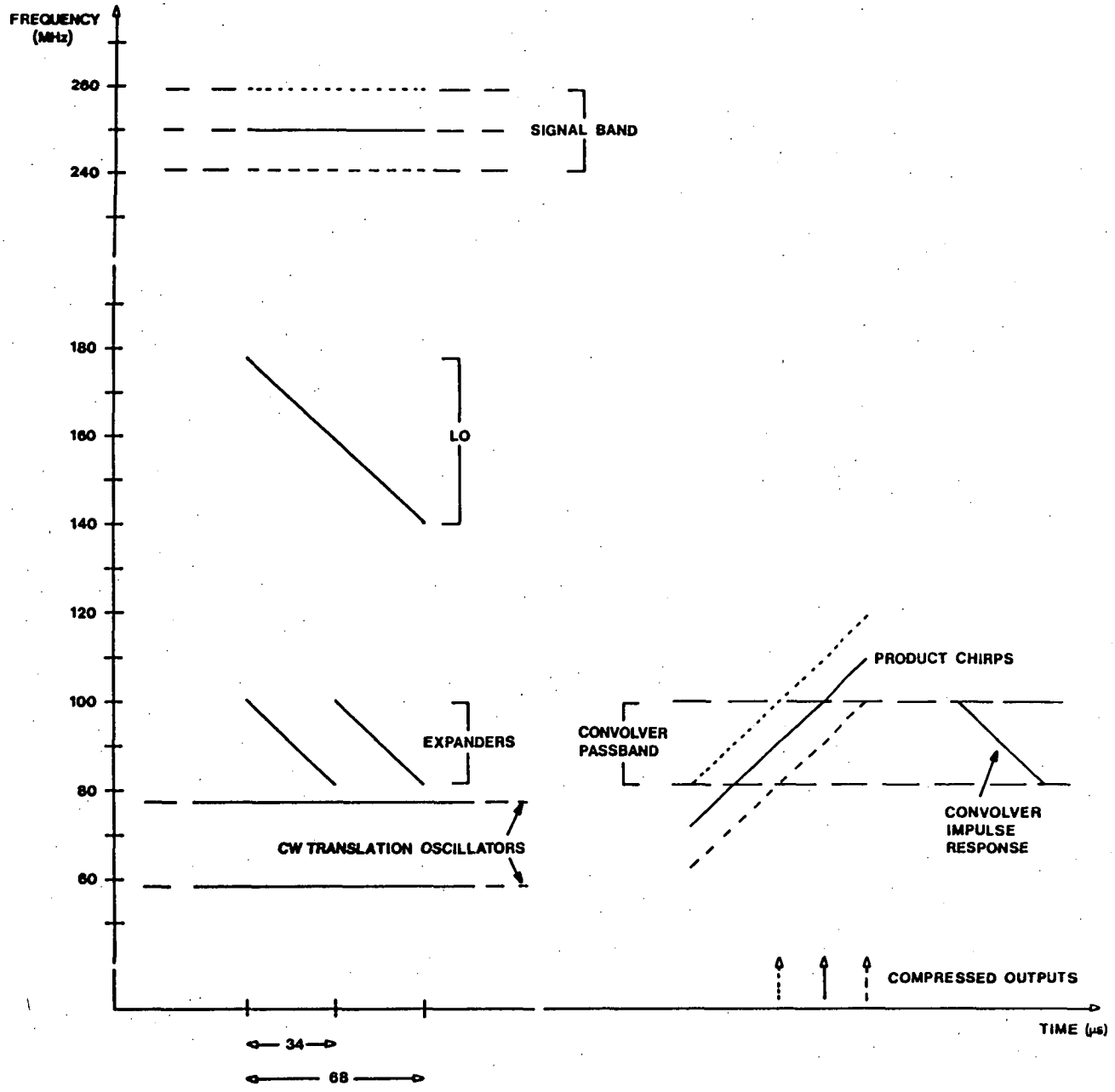


Fig.6-6: Frequency-time diagram for the compressive receiver discussed in the text. The product chirps from the signal mixer are shifted along the time axis for clarity.

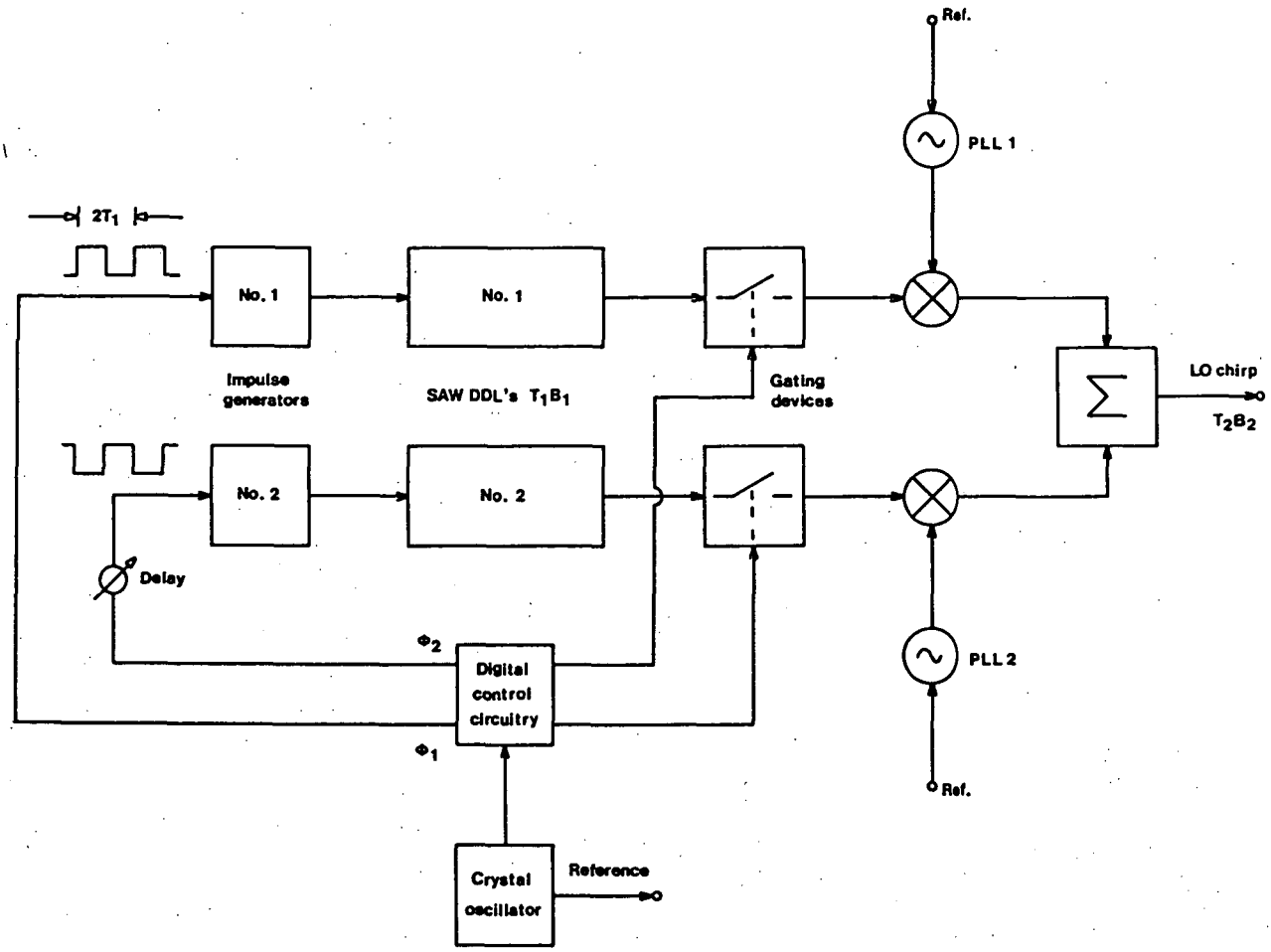


Fig.6-7: Active method of producing a double duration, double bandwidth LO chirp.

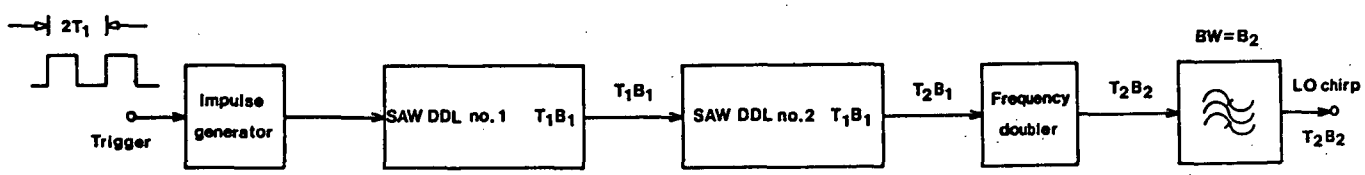


Fig.6-8: Passive method of LO chirp generation.

bandwidth product of four times the constituent chirps (Signal Technology, 1980). Although simpler than the previous method, the scheme has the disadvantage of producing an LO signal with a rather large amount of Fresnel ripple due to the chirp-chirp convolution in the second expander DDL. Amplitude ripple can be removed by limiting, although the LO noise floor is then raised. Phase ripple is not removable. Nevertheless, for many applications the method produces acceptable results.

Extension of the active LO generator to 100% duty cycle mode is possible and involves impulsing the DDL's at twice the frequency used for 50% duty cycle. Depending on the actual approach adopted, high duty cycle operation may involve extra electronic gating at the DDL outputs. In some cases overlapping chirps need to be produced from a single DDL. Since the device is linear this causes no problem and, in a linear LO chain, overlapping chirps can be amplified and frequency translated with minimal intermodulation. The non-linear doubler in the passive generator precludes its use at 100% duty cycle.

Several comments can be made concerning more unusual chirp generator configurations. SAW DDL's are expensive compared with digital control circuitry so it is often worth investigating the possibility of generating an LO chirp by impulsing a single DDL at intervals equal to its chirp duration and then switching the output chirp between frequency translation stages. The feasibility of such schemes depends upon the convolver and expander absolute delays and upon the analyser duty cycle. For given parameters the risk of feeding spurious signals into the signal mixer can be assessed. Campbell et al. (1981) describe a 100% duty cycle compressive receiver in which one expansion DDL, elaborate digital control circuitry and extensive LO switching is used to produce a workable spectrum analyser.

To ensure that compressed pulse sidelobe specifications can be met, amplifiers and other devices in the chirp generator must produce a minimum of phase and amplitude ripple over the chirp bandwidth and must, in addition, exhibit good transient response so that the "tone-burst" chirp signals are not distorted. The final high-level amplifier must be capable of delivering the required output power (~ + 10 dBm) without compression and, to ensure that the LO noise floor is determined

primarily by thermal noise at the dispersive filter outputs, small-signal stages must use low-noise amplifiers. All of the amplifier specifications listed are met by many of the wideband, open loop hybrid modules now available. Most modules offer the advantage of a well-defined input and output impedance, making them suitable as sources and loads for mismatch sensitive devices such as DDL's, hybrids, filters and mixers.

6.6 INPUT, SIGNAL MIXING AND CONVOLUTION STAGES

The design of these three stages is straightforward and some aspects have already been referred to in earlier sections. The analyser input should be band-limited to prevent saturation of the signal amplifier by out-of-band signals or by broadband noise. Band-limiting also reduces the level of uncompressed image chirps at the convolver output. The input amplifier can be a wideband, high-level modular device of the type mentioned in the previous section.

The signal mixer should be a double balanced type (DBM) to ensure maximum isolation between ports. It is essential that the mixer be properly terminated at all ports. Experimental investigations have demonstrated the value of using mixers having operational bandwidths much larger than the nominal bandwidth required. Deterioration in the balance and distortion performance of typical mixers at high frequencies is sufficient to degrade the spectrum analyser performance if marginally rated devices are used. If a suitable amplifier is available to raise the LO level to about +17 dBm, there is some advantage in using a high-level mixer since distortion is reduced.

If the image band rejection of the input filter is not sufficient to cause uncompressed image chirps at the analyser output to fall below the noise floor, it is desirable to use an image rejecting (single-sideband) mixer. Where the signal amplifier is a broadband type, there is always a slight gain in dynamic range if an SSB mixer is used since the noise contribution of the amplifier is halved, lowering the output noise floor by 3 dB if the noise floor is set by the signal path (section 6.3).

When the compressive receiver incorporates an input bandpass filter ($B_{filter} = B_1$) to allow operation at 100% duty cycle, the signal mixer should be a bi-linear type so that linear superposition of two

chirps offset in frequency occurs at the L port (figure 4-8). The wideband analog multiplier described in the next section may be suitable in this role.

In a coherent SAW processor (figure 4-10) the extensive output splitting means that the post-convolver amplifier must be a high-level type (typically +30 dBm output) to ensure good video dynamic range. Broadband, modular amplifiers meeting this requirement are available commercially.

6.7 COHERENT ANALYSER POST-TRANSFORMATION PROCESSING

The post-convolver splitting shown in figure 4-10 can be achieved with broadband in-phase power dividers and quadrature hybrids. The splitting network must function properly at the convolver mid-frequency and over a frequency range equal to the compressive receiver output bandwidth. The required network function can be obtained by choosing appropriately specified hybrids and splitters, correctly terminating all devices, and minimising unwanted differential phase shifts in correlator signal paths.

Figure 6-9 shows a practical distribution system and illustrates the design principles involved. In this arrangement the level to each detector or correlator is the same for all four processor channels because level equalising attenuators are included in the detector signal paths. Identical quadrature hybrids are used in a scheme designed to compensate for the insertion and transmission phase characteristics of the devices. A resistive two-way divider splits the in-phase output of the channel 1 hybrid and two similar networks are used at the outputs of the channel 2 hybrid. This arrangement results in a distribution network with little unwanted phase shift.

The problem of direct high frequency correlation of wideband SAW compressive receivers is a challenge to existing technology. In a direct correlation approach the major requirement is for a four-quadrant, bi-linear, analog multiplier with input and output bandwidths of at least the same order as the SAW spectrum analyser output bandwidth. To be comparable with good diode square-law detectors the device should have a usable dynamic range of around 30 dB. It must also exhibit a low level of self-detection (self-squaring) and feed-through from each

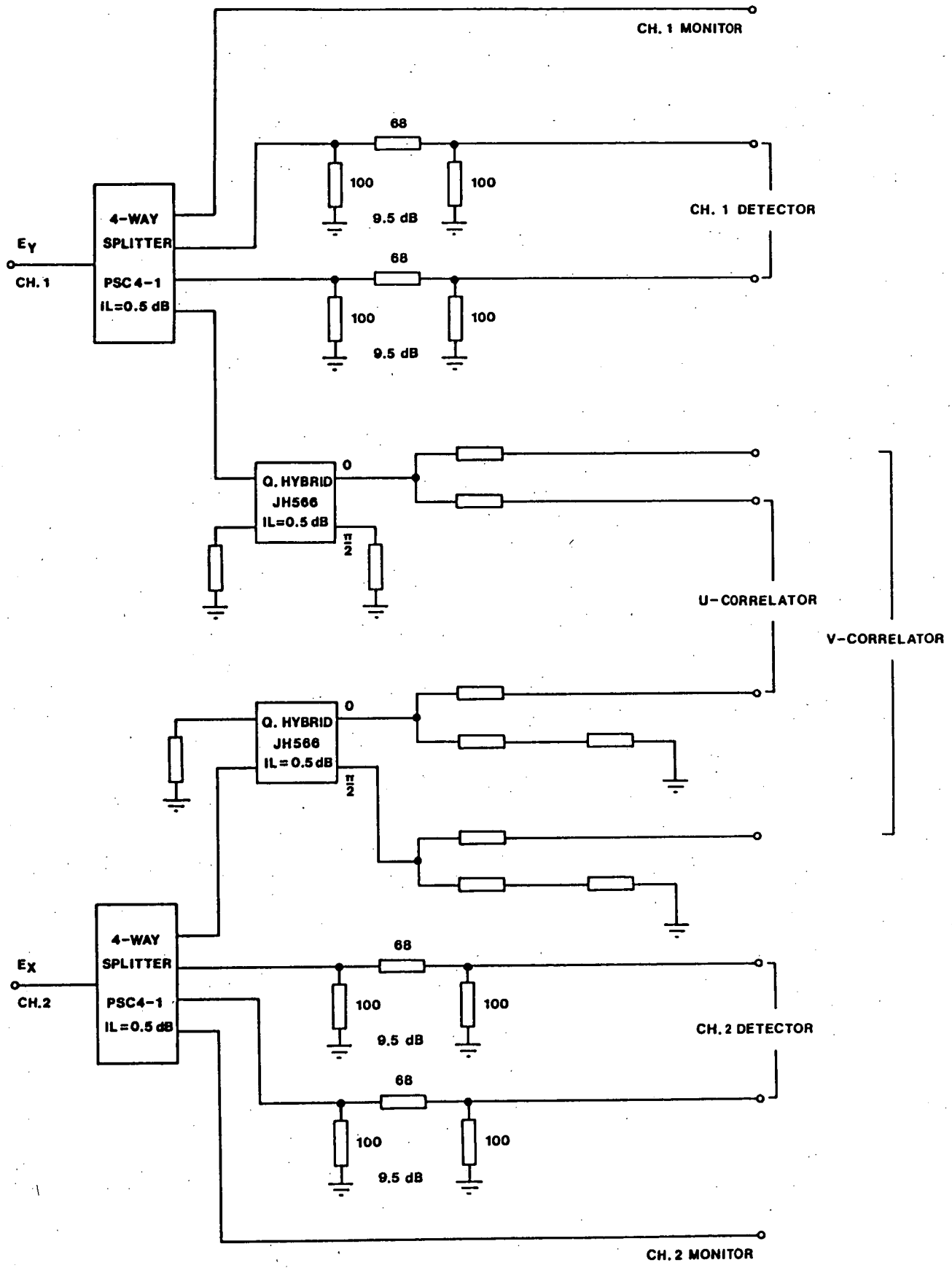


Fig.6-9: Post-convolver signal splitting network. The components shown are suitable for 60 ± 15 MHz operation. The four-way splitter is made by Mini-circuits Laboratories and the quadrature hybrid is an Anzac type. Unmarked resistors are 50Ω. The network input and output impedances are 50Ω.

input and should maintain a flat amplitude and linear phase response across the analyser output bandwidth.

Design of post-transformation stages is simplified if the multiplier accepts signals at the convolver centre frequency. The convolver output band can be translated down to ease the input frequency requirement but it is desirable that the multiplier input band be well-separated from the video output to minimise the effect of feed-through. In most cases this means that a VHF input capability is necessary, regardless of whether a post-convolver frequency translator is used.

The need for accurate high frequency operation excludes several multiplier options available to designers of electronic systems operating at lower frequencies. For example, neither Hall effect nor magnetostrictive multipliers are suitable for VHF use. Diode ring (double balanced) mixers function into the UHF range but can be rejected because of very limited dynamic range when used as bi-linear multipliers (Wilcox, 1959). Colvin et al. (1979) describe a prototype bi-linear mixer based on non-linear wave interaction in a SAW device. The device has a 60 dB dynamic range with single tone inputs but its effectiveness on broadband signals has yet to be demonstrated. The mixer seems most suitable in applications where a narrowband RF output is acceptable.

A transconductance multiplier is the most promising type if good high frequency operation is required, although wideband operation presents difficulties. The analyses of Frater (1964) and Gilbert (1968) show that accurate active device matching is essential for satisfactory performance. A high performance multiplier must normally be fabricated as an integrated circuit to obtain adequate matching and to minimise parasitic effects. Even so, fine adjustment of internal currents is always necessary to optimise performance.

Frater and Gough (1981) have demonstrated a non-integrated multiplier having a 1 GHz input bandwidth but a small output bandwidth. Their multiplier was designed using computer modelling techniques and required careful trimming for correct operation. Intensive military research has reportedly produced a workable high frequency multiplier with wide input and output bandwidths (Estrick and Judd, 1982). Available information on the device has been summarized in section 4.7.

Tests of common IC multipliers such as the Motorola MC1495 and MC1496 in a SAW spectrometer correlation role indicate that the main problems are limited input frequency range, excessive feed-through and self-squaring at high input frequencies, excessive internal phase error, and a limited output voltage swing when the devices are optimised for operation with a wide output bandwidth. The output level restriction, as well as a slew rate limitation, means that several stages of wideband amplification are necessary to raise the final output to the 1 V peak-to-peak figure required by typical high speed A/D converters. With low output signal swings, internal IC noise and the noise contribution from video amplifiers is sufficient to restrict correlator dynamic range to less than about 12 dB. In addition to exhibiting a poor signal to thermal noise ratio, prototype multipliers showed drift and flicker noise comparable with typical high frequency output swings. This prevented the construction of a wideband direct-coupled multiplier.

Figure 6-10 is the circuit diagram of a new VHF multiplier designed for use as a correlator and square-law detector in SAW spectrum analysers. The component values shown are suitable for the prototype coherent processor described in chapter seven. This processor has an output bandwidth of 20 MHz and a convolution centre frequency of 60 MHz. Many of the multiplier performance figures referred to in the following description are derived from the results of tests reported in section 7.6. Oscillograms and graphs are also included in that section.

The correlator module is based on a low-cost integrated multiplier (Analog Systems MM109) designed for use in analog graphics processing applications. The specified IC input bandwidth is 75 MHz and the output bandwidth is d.c.-55 MHz. The multiplier output bias level is +3.6 V and in the basic circuit the output is a.c. coupled to a post-correlation amplifier. Direct coupling is also feasible and is discussed shortly.

The X-input delay line compensates for internal phase shift and the input offset potentiometers are included to minimise self-squaring. To adjust the X-input control, the X-input is terminated and the control adjusted for minimum output due to a signal at the Y-input. A similar procedure is used to set the remaining control. With correct adjustment, the output component due to input self-detection is about 1% of the wanted product.

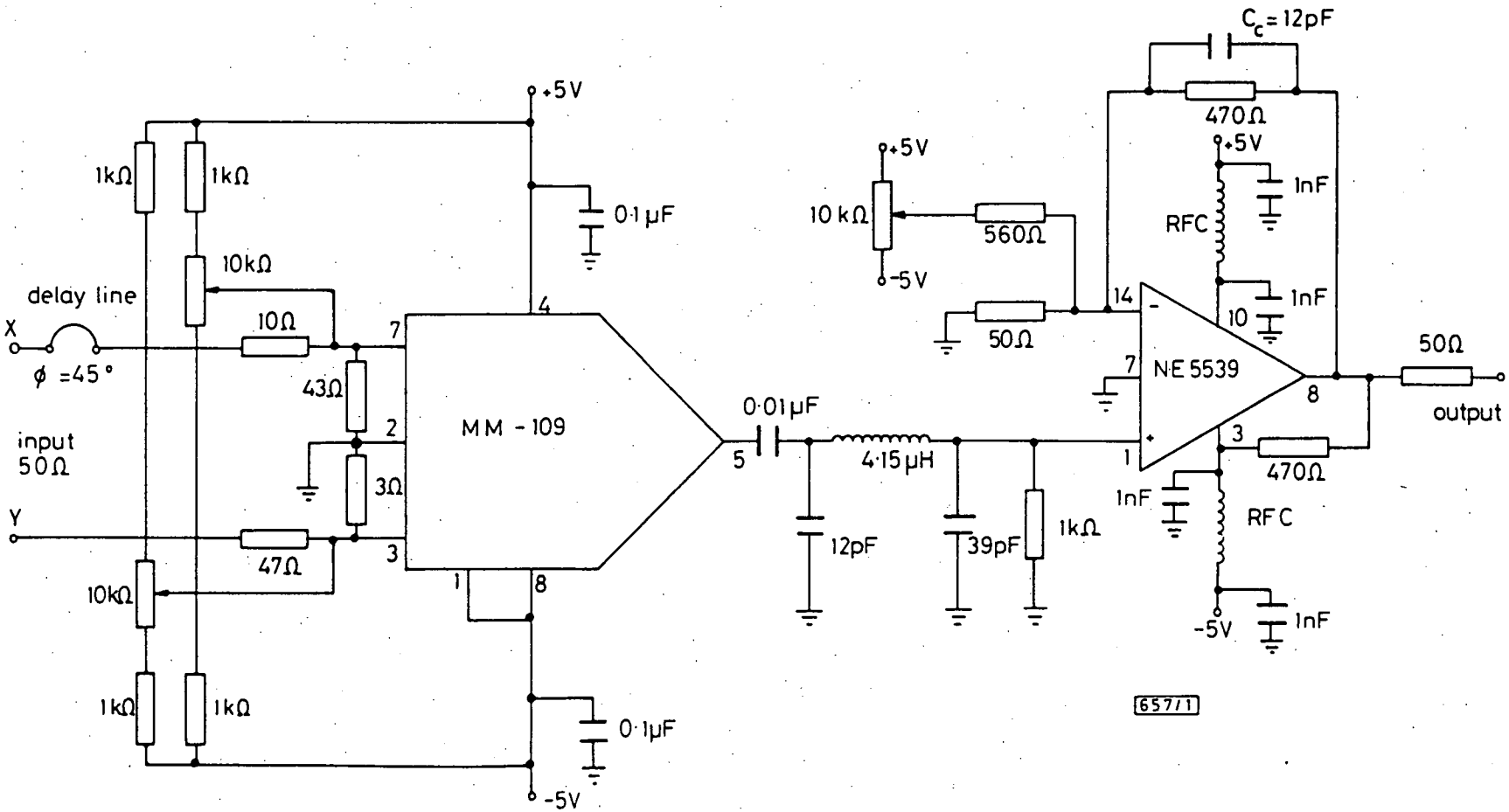


Fig.6-10: Basic VHF correlator using an MM109 multiplier IC. The component values shown are for 60 MHz operation and give a -3 dB video bandwidth of 20 MHz (the NE5539 stage bandwidth is 25 MHz). The maximum RF input is ±0.5 V; the corresponding output range is ±1 V around the mean level.

The post-correlation filter is a third-order Butterworth type having an input impedance designed to match the 350 ohm output impedance of the multiplier IC. See section 4.6 for further comments concerning the choice of filter. The non-inverting gain stage uses a very wideband operational amplifier (Signetics NE5539) with a gain-bandwidth product of 1.2 GHz, a slew rate of 600 V/μs, and a full-power bandwidth of 48 MHz. Internal design restricts the NE5539 output swing to ±2 V. The voltage gain is 21 dB and C_c is chosen to yield a frequency response which is 3 dB down at 25 MHz.

The amplifier input impedance is 1 kΩ, set by the shunt resistor at the non-inverting input. This value applies minimal loading to the multiplier IC output and preserves most of the available voltage swing. At the same time, 1 kΩ is around the highest filter termination resistance which allows reliable HF filter realisation. The feedback resistors at the inverting input have low values to minimise the effect of parasitic admittances around the circuit.

The correlator module functions as a detector at input frequencies of several hundred megahertz and investigations suggest that reasonably accurate correlation can be obtained at such frequencies. In applications where the input frequency band is much higher than the video bandwidth, the post-correlation filter can be simplified or omitted, leaving the amplifier to function as a lowpass element. With optimum compensation (C_c ~ 1.6 pF) the amplifier response exceeds 50 MHz with the component values given, so full use can be made of the MM109 output bandwidth where necessary.

The circuit of figure 6-10 was developed originally to allow assessment of the MM109 performance using high SNR spectra viewed on an oscilloscope. In astronomical applications direct-coupled correlators or detectors are necessary to preserve mean output levels and to prevent waveform distortion of transient or impulsive signals. Figure 6-11 shows a direct-coupled version of the new correlator. Direct coupling is achieved using an MM109 on-chip buffer and an external passive level-shift network. The correlator module also incorporates a second video amplifier, a clipper and a buffer amplifier. Clipping levels restrict the output to -0.5 ±0.5 V in the prototype radiometer. The buffer amplifier gives the module the capability of

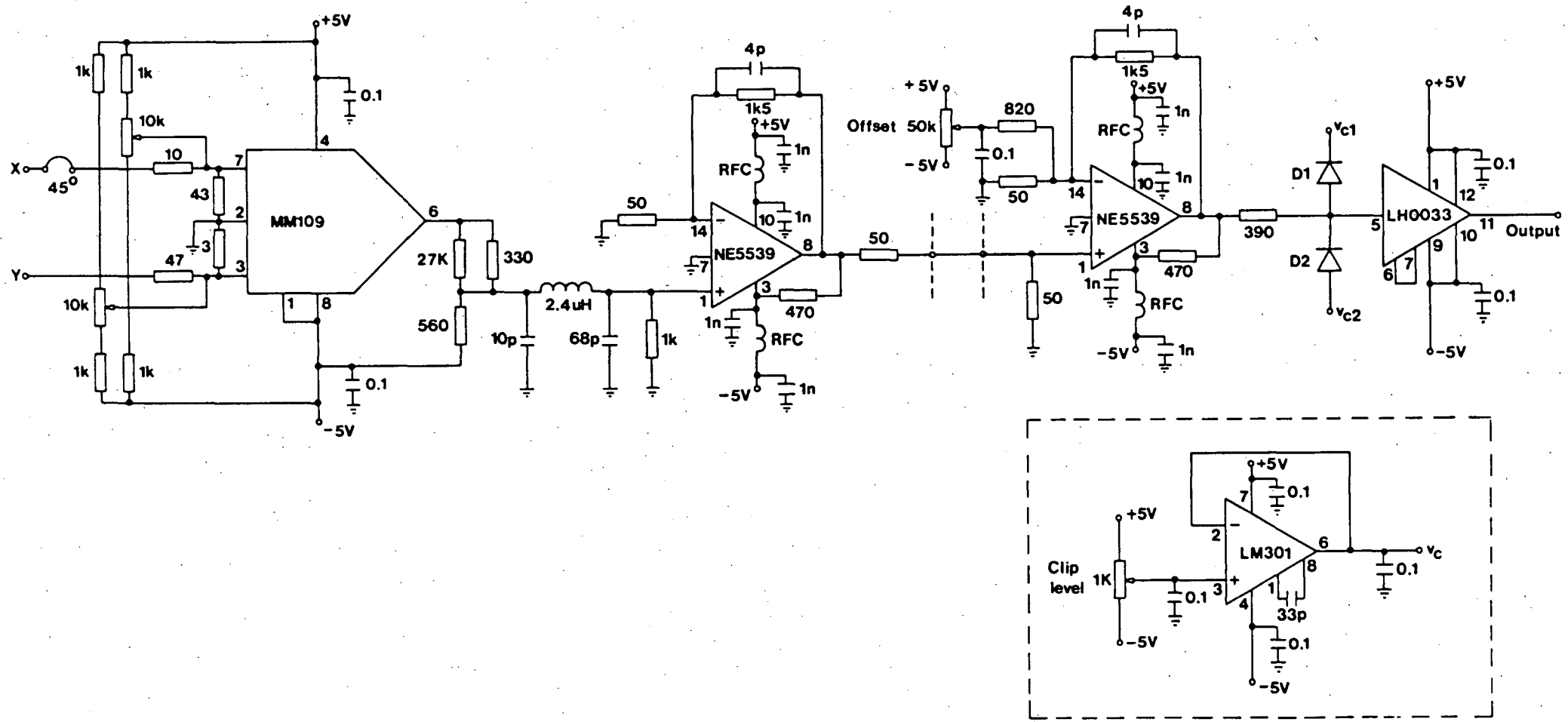


Fig.6-11: Direct-coupled VHF correlator module. The component values shown are for 60 MHz operation and a video bandwidth of 20 MHz. The input impedance is 50Ω and the module is capable of driving 50Ω coaxial lines. Diodes D1 and D2 are Schottky types; the circuit used to derive the clipping reference levels is shown in the inset. By-pass capacitors are low inductance types. The construction style is described in the text.

driving 50 ohm coaxial cable without the video pulse extension which would occur if the output were taken directly from the clipper.

The use of an extra video stage and a buffer amplifier, together with a gain increase in the first video amplifier, leads to lower-frequency break-points and a more rapid roll-off in the video frequency response. A Gaussian post-multiplier filter is therefore substituted for the Butterworth type shown in figure 6-10. This improves the time response considerably (section 7.6).

As with any high frequency circuit, layout and de-coupling are critical for correct operation. The MM109 and first video stages of the four prototype direct-coupled multiplier modules are constructed on two identical double-sided printed circuit boards, one side of which is a continuous ground-plane. One PCB holds the two correlators, the other the square-law detectors. The style of construction is shown in figure 7-4(b). With the exception of the phase compensation line, all signal splitting components after the four-way post-convolver divider are contained on the PCB's. The delay line is a differential length of coaxial cable in the leads connecting the post-convolver splitter to the detector or correlator enclosures. The remaining video circuitry for the four multipliers is constructed using the same PCB techniques and is contained in a separate enclosure.

The MM109 internal shield (pin 1) is connected to the negative supply rail and the IC is by-passed directly at the power pins by low inductance ceramic or chip capacitors connected to the ground-plane. The de-coupling networks in NE5539 stages are constructed similarly with the chokes being F43 ferrite beads slipped over wire links in the PCB supply rail tracks. Unused pins on the NE5539 DIL packages are returned directly to the ground-plane.

The relatively high output capability of the MM109 minimises the contribution of drift and flicker processes at the direct-coupled module output. Even so, a warm-up period of about one hour is necessary for un-attended operation of the equipment. Thereafter the drift is slow with a quasi-periodic drift cycle of two hours being typical. After warm-up the output remains within ±50 mV of the nominal offset. These figures apply in an ordinary air-conditioned environment. Slow drift is acceptable in the observation of dynamic cosmic emission because

both signal and baseline move simultaneously, allowing calibration methods to remove the effect. For continuum source observations, radiometer configurations designed to minimise the effect of gain variations and drift could be implemented (Tiuri, 1964) although an auto-offsetting system could also be used.

The results of the tests described in section 7.6 indicate that the useful dynamic range of the basic correlator (figure 6-10) when used as a square-law detector is about 28 dB, being limited by the noise floor of the multiplier IC. The range between the point of 5% departure from linearity and the noise floor is 25 dB. Equivalent input noise voltages at the MM109 inputs are given as 18 nV/$\sqrt{Hz}$ and 25 nV/$\sqrt{Hz}$. Assuming independence of the noise contributions, the resulting 31 nV/$\sqrt{Hz}$ taken over approximately 30 MHz equivalent video noise bandwidth gives about 0.2 mV r.m.s. at the IC output, or around 2 mV at the NE5539 output in figure 6-10 if the contribution of the low-noise operational amplifier is small. The measured noise voltage is around 1.5 mV, identified as due largely to the MM109. Since the predicted and measured levels are in good agreement, it is unlikely that the correlator noise floor can be reduced further.

In the direct-coupled multiplier the r.m.s. noise level at the module output is about 50 mV. The maximum output range is 1 V so the dynamic range is restricted to 13 dB assuming square-law detection is the required operation. In the prototype radiometer the A/D converters are four-bit types with an input range of 1V. On the basis of 3 dB/bit quantization dynamic range (section 5.1), the A/D converter remains the principal limitation on dynamic range when the direct-coupled detector is used. Given the signal level obtainable from the SAW processor and the output limitations of the MM109, the second video stage in the direct-coupled module ensures sufficiently high output levels for effective quantization of noisy spectra. The clipper prevents saturation of the A/D converter analog stages by noise peaks. Tests conducted using the prototype system show that the clipper is an important practical refinement.

CHAPTER 6 - REFERENCES

Campbell, P.A., Grant, P.M. and Arthur, J.W.: Extension of time-bandwidth product in SAW spectrum analysers. Electron. Lett., Vol.17, No.16, August 1981, pp.576-578.

Coekin, J.A.: "High speed pulse techniques". Ch.5, Pergamon, 1975.

Colvin, R.D., Carr, P.H. and Charlson, E.J.: SAW bi-linear mixer. Proc. 1979 IEEE Ultrasonics Symp., pp.721-724.

Davie, M.C. and Arthur, J.W.: A comparative study of SAW compressive receivers for radar intercept. Racal-MESL Microwave Ltd., Internal Report, Edinburgh, 1979.

Estrick, V.H. and Judd, G.W.: A SAW analog correlator using the chirp transform. Microwave Journal, Vol.25, No.7, July 1982, pp.97-106.

Frater, R.H.: Accurate wideband multiplier square-law detector - Rev. Sci. Instrum., Vol.35, No.7, July 1964, pp.810-813.

Frater, R.H. and Gough, R.G.: Wide band multiplier correlator. Rev. Sci. Instrum., Vol.52, No.5, May 1981, pp.724-726.

Gilbert, B.: A precise four-quadrant multiplier with subnanosecond response. IEEE J. Solid State Circ., Vol.SC-3, No.4, Dec. 1968, pp.365-373.

Kowatsch, M. and Lafferl, J.: Generation of nanosecond high voltage pulses with high repetition rate. Electron. Lett., Vol.16, No.5, Feb.1980, pp.196-198.

Longhurst, R.S.: "Geometrical and physical optics". Ch.12, Longman, 1973.

Signal Technology Ltd.: SAW spectrum analysis. Application note, 1980.

Tiuri, M.: Radio astronomy receivers. IEEE Trans. Ant. Prop., Vol.AP-12, Dec. 1964, pp.930-938.

Wilcox, R.H.: Use of a diode ring as a four quadrant multiplier. Rev. Sci. Instrum., Vol.30, No.11, Nov. 1959, pp.1009-1011.

CHAPTER 7 A SAW COHERENT PROCESSOR

7.1 INTRODUCTION

In this chapter a prototype SAW coherent processor (designed as a Stokes polarimeter) is described and the results of laboratory performance tests are reported. The instrument, shown in block diagram form in figure 7-1, is a practical implementation of the processor illustrated in figure 4-10 and described in Section 4.7. Design principles for major modules have been discussed in chapter six and, in particular, the post-convolution circuitry is detailed in section 6.7.

7.2 SAW DISPERSIVE DELAY LINE CHARACTERISTICS

The two SAW compressive receivers are based on low-cost, radar DDL's having complementary chirp slopes (Signal Technology models DS1301/1302). The nominal device chirp time (T) is 2.2 μs and the specified bandwidth (B) is 25 MHz (centred at 60 MHz), giving TB = 55. The absolute delay for both devices is 1.4 μs measured from the time of application of an impulse to the beginning of the chirp output. The measured chirp duration is 2.4 μs and the bandwidth is close to 30 MHz, giving k = 12.5 MHz/μs and TB = 72.

The convolver is weighted to reduce compressed pulse sidelobes, the first sidelobe level being given as -32 dB relative to the peak. For TB = 72 the Fresnel or gating sidelobes around t = ± T/2 are approximately 40 dB down. The weighting function broadens the compressed pulse to a width of 50 ns, resulting in a pulse bandwidth of about 20 MHz and a pulse broadening factor of 1.5.

The substrate material for the DDL's is $LiNbO_3$ and in common with most SAW devices the insertion losses are fairly high: 36 dB for the expander and 28 dB for the compressor. The devices are mounted in an hermetically sealed DIL header measuring 13 mm × 25 mm.

Enquiries to the SAW device designers confirmed suspicion that the DS1302 convolver DDL is reciprocal ripple compensated, a fact not stated on the manufacturer's data sheets. Although compensated convolvers are undesirable in a spectrum analyser (section 6.2), the low cost of the DS1301/1302 devices relative to more suitable DDL's was sufficient incentive to develop a prototype instrument using the DS1301/1302 combination.

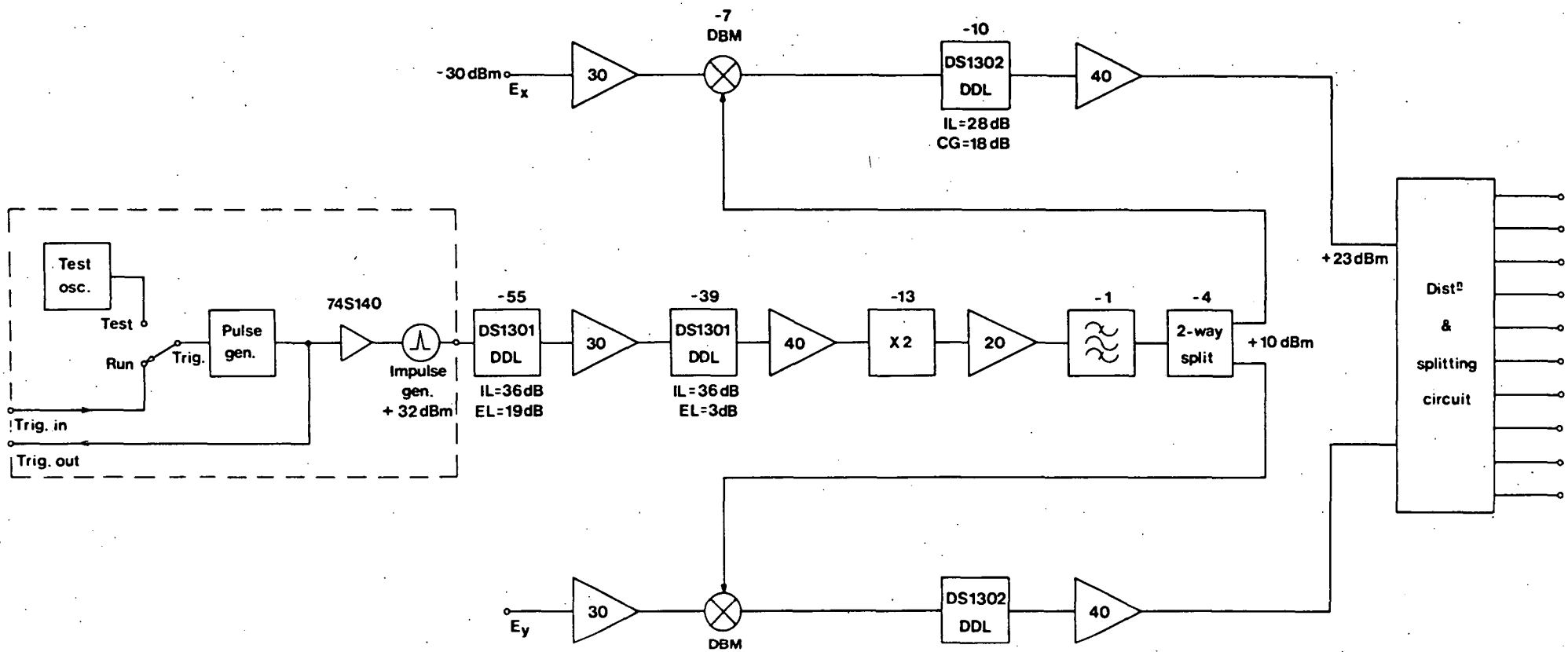


Fig.7-1: Block diagram of the prototype SAW coherent processor. Gain and loss figures are in decibels and the signal path levels quoted are approximate maximum values for a CW input. Details of the post-transformation splitting network are shown in figure 6-9.

To verify the feasibility of constructing a coherent SAW processor, the amplitude and phase responses of the two SAW convolvers were compared using an HP 8754A network analyser and an HP 8405A vector voltmeter. Over the entire DDL passband the amplitude responses remained matched to with ±0.5 dB. The absolute phase difference at 60 MHz was 30^{o}. The differential phase response was linear to within $\pm 1^{o}$ over the range 50-65 MHz and over the range 45-75 MHz the departure did not exceed $\pm 3^{o}$. The excellent matching of two production line SAW devices was cause for optimism about the performance of a coherent analyser.

7.3 PROCESSOR DESIGN

Compressive receivers using both active and passive chirp generators (section 6.5) were tested and the performances compared. In both cases, analysis of the compressed pulse spectrum showed increased ripple amplitude and periodicity as a CW input signal was tuned towards the edges of the spectrum analyser passband. As expected, the ripple was manifested in the time domain by increased sidelobe levels for signals at the passband edges. The near-sidelobe level of a compressed pulse corresponding to an input at the analyser passband centre was -22 dB (-44 dB after square-law detection), about 10 dB higher than that measured in a simple compression loop. The high sidelobe level at centre frequency and the strong dependence of sidelobe level on relative input frequency was attributed to mismatch between the Fresnel characteristics of the pre-convolution chirp and the chirp impulse response of the ripple compensated convolver.

Since the sidelobe performance of an instrument using the DS1301/1302 DDL's is limited by the convolver ripple compensation, a spectrometer design using the passive LO chirp generation method (figure 6-8) was implemented with the expectation that the disadvantages associated with this scheme would be outweighed by the convolver limitations. This supposition is confirmed in practice.

Using actual DDL parameters, the predicted bandwidth of the prototype spectrometer is 30 MHz with a frequency resolution cell width of $1.5 \times 1/(2.4\,\mu s)$ or 630 kHz, resulting in 48 frequency resolution cells over the passband. An input bandpass filter and a bi-linear signal mixer are not used, so the maximum duty cycle of the processor is nominally 50% (4.8 μs impulse pulse repetition interval).

The timing of the prototype compressive receivers is illustrated in figure 7-2 and the idealized frequency response is also shown. In practice, chirp signals falling only partially within the convolver result in outputs considerably below the level implied in figure 7-2 because of the tapered convolver frequency response. This allows a duty cycle of 60% (4.0 μs impulse PRI) before the full-compression data window is corrupted by aliased components.

A frequency-time diagram for the prototype spectrum analysers is shown in figure 7-3. The input band cannot be displaced from the convolver passband if a passive LO generator is used in conjunction with expansion and convolution SAW devices having the same centre frequency and complementary chirp slopes. However, investigations showed that mixer breakthrough is minimal, being just above the noise floor set by other factors. The inputs to the coherent IF processor are assumed to be lowpass filtered in stages preceding the instrument, although the apparent problem with out-of-band signals above 90 MHz does not occur in practice because of strong suppression of chirps from signals in this range by the weighted convolver.

The thermal sensitivity of the processor has been estimated using the methods described in section 4.5. At the passband centre the predicted sensitivity is 1.2 ns/^{o}C, rising to 1.5 ns/^{o}C at the passband upper extremity. Passive temperature stabilization methods should be satisfactory since the worst-case sensitivity amounts to only 0.03 resolution cell/^{o}C.

The impulse generator is an avalanche type (figure 6-3) and is triggerable from either an internal test oscillator or an external source. A trigger output provides convenient triggering when viewing the output spectrum on a CRO. The impulse generator transistor is driven from a 74S140 line driver IC to ensure a clean triggering characteristic.

TRW CA2820 broadband amplifiers are used throughout the processor. These hybrid modular amplifiers have a gain of 30 dB over a bandwidth of 500 MHz, a third-order intercept point of +39 dBm and a noise figure of 5.5 dB when used with an 18 V supply. Only one amplifier, three RF chokes and two de-coupling capacitors are required to implement a complete gain block. Amplifier stages using the hybrid modules exhibit an excellent transient response and are ideal for use with chirp

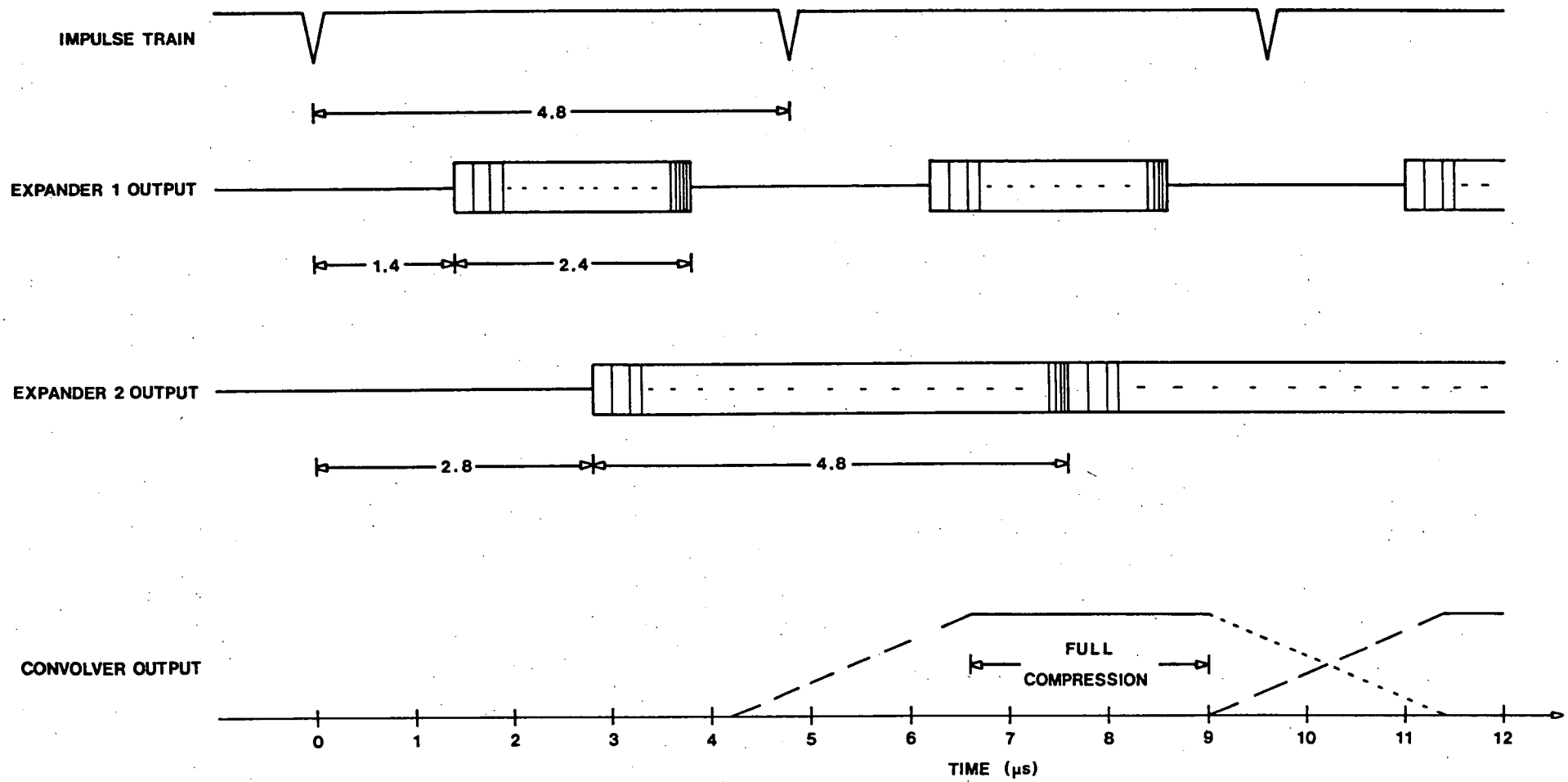


Fig.7-2: Idealized timing diagram for the prototype compressive receivers operating at 50% duty cycle.

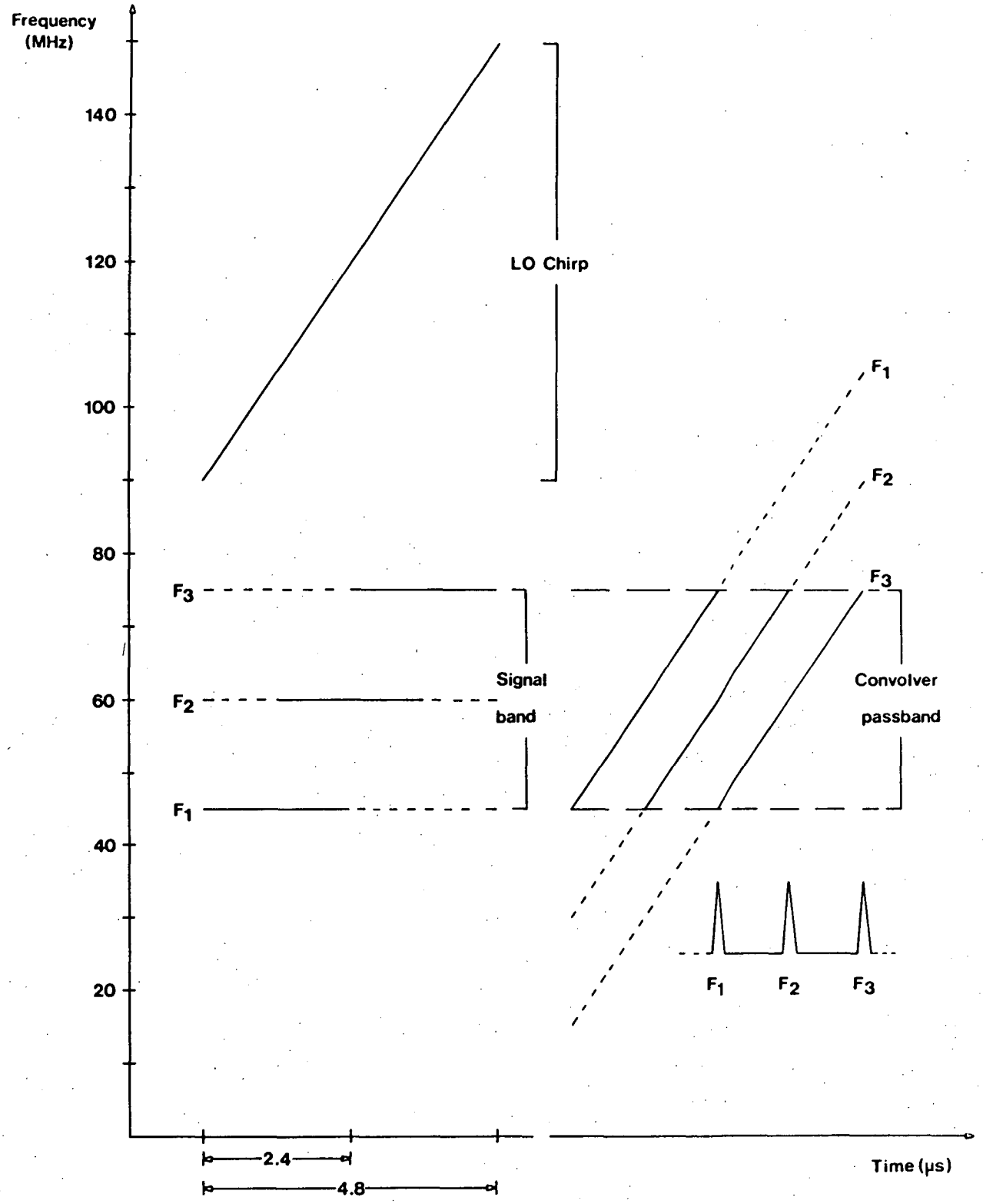


Fig.7-3: Frequency-time diagram for the compressive receivers. The product chirps are shifted along the time axis for clarity since the passband of the down-chirp convolver covers the same frequency range as the signal band. The relative positions in time of the compressed outputs are also shown (the convolver absolute delay is neglected). The spectrum of the compressed signals is centred at 60 MHz.

signals. Where gains other than 30 dB are indicated in figure 7-1, CA2820 amplifiers and resistive pads are used to obtain the required values.

7.4 CONSTRUCTION

In assembling the processor, care was taken to minimise coupling between the two channels and between high-level stages (such as the impulse generator) and low-level sections. The impulse generator, LO chirp generator, signal IF stages and post-transformation circuitry are enclosed in separate brass enclosures. The four SAW devices are mounted in a common enclosure. Double sided printed circuit board construction is used in most sections of the instrument. One side of the PCB's is a continuous ground-plane and is connected to the brass enclosures via low impedance paths. PCB construction facilitates easy phase matching of the two signal channels and is mandatory if coupling and instability problems are to be avoided. All low-voltage power leads entering enclosures are by-passed with 0.1 μF ceramic capacitors outside the enclosure, 1000 pF feed-through capacitors and internal 10 μF tantalum electrolytic capacitors. Extensive use is made of LC de-coupling networks between stages contained in a common enclosure.

The prevention of stray coupling of the excitation impulse is difficult and thorough power lead by-passing proved to be essential from this consideration alone. In the impulse generator enclosure, small-value capacitors having a low series inductance are used in addition to the usual by-pass components. To further minimise stray impulse coupling, the pulse bandwidth is restricted deliberately by a 100 MHz lowpass filter (third-order Butterworth). The filter has a negligible effect on impulse components around 60 MHz but the high frequency roll-off is very effective in removing residual coupling to the sensitive correlators and detectors.

In the DDL enclosure, the four SAW devices are PCB-mounted and the device cases bonded thermally to a brass plate. This mounting method minimises differential temperature changes between the DDL's. Extensive internal shielding and grounding is used within the enclosure to prevent coupling between the devices.

No special effort has been made to minimise the size of the prototype instrument but even so, the entire SAW processor is contained in a rack-mounting box 430 mm × 265 mm × 400 mm. A separate rack-mounting enclosure contains the instrument power supplies as well as most of the power supplies for the digital spectrum integrators described in the next chapter.

Figure 7-4(a) is a photograph of the processor and power supply. The eight front-panel bulkhead BNC connectors carry the following signals:

(a) two 60 MHz IF inputs;
(b) two correlator outputs, two intensity outputs;
(c) an impulse generator trigger input and a CRO trigger output (both TTL compatible).

A multi-pin power connector at the rear of the unit is the entry point for all three processor power supply rails. These are the:

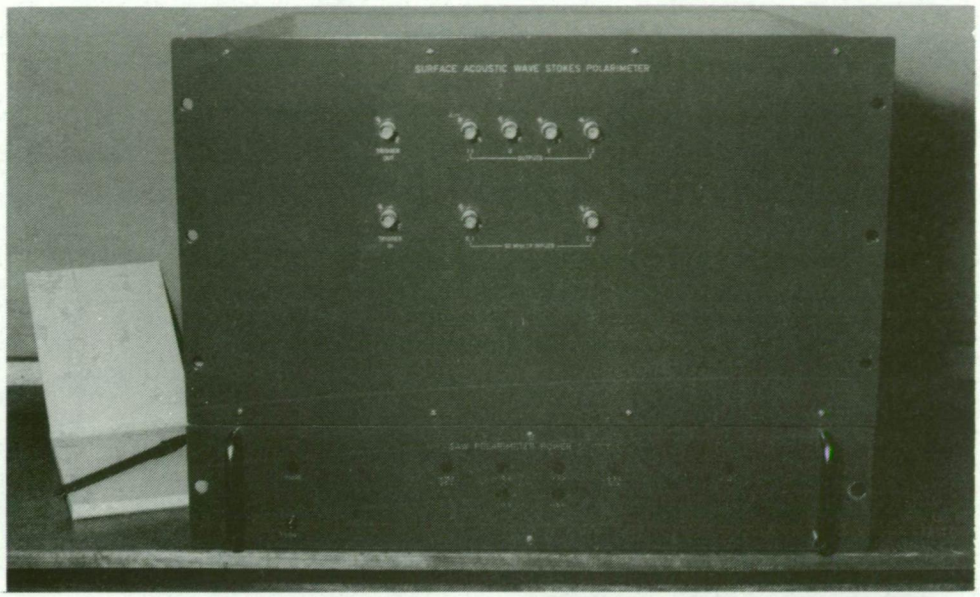
(a) +125 V regulated impulse generator supply (0.03 A);
(b) +18 V regulated hybrid RF amplifier and general supply (2.5 A);
(c) ±15 V regulated detector and correlator supply (1 A).

On-board three-terminal regulators provide further voltage reduction and stabilization where required. Figure 7-4(b) is an internal view of the instrument with the covers removed from the detector and correlator enclosures.

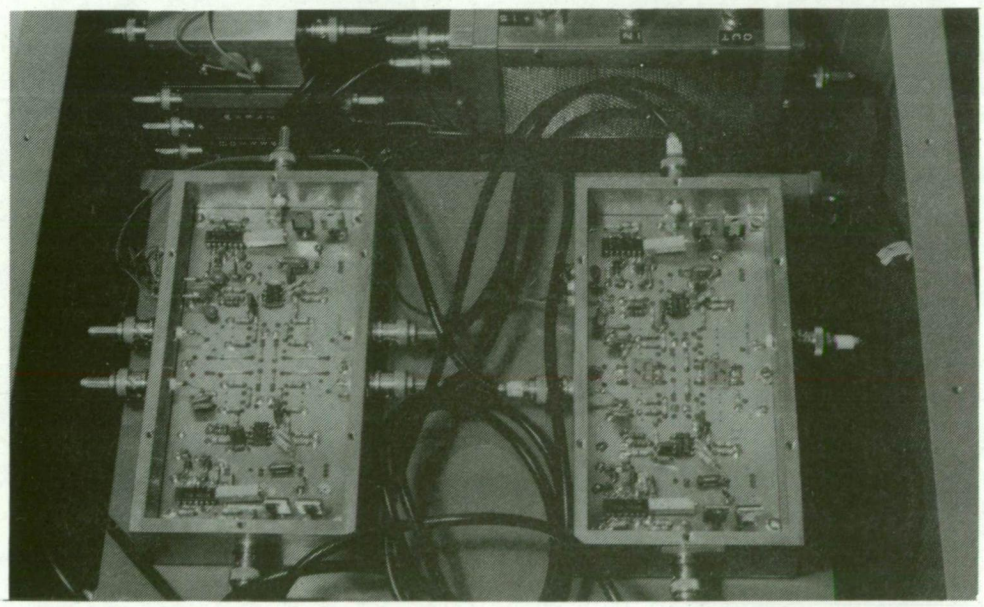
7.5 ALIGNMENT

Alignment of the SAW processor presented no difficulty since the use of broadband hybrid amplifiers and PCB construction limited greatly the possible sources of amplitude and phase error. No amplitude correction was necessary since the two compressive receiver passbands were matched to within a fraction of a decibel (section 7.6).

Phase compensation commenced with the adjustment of the correlator delay lines (figure 6-11). The two input lines leading to the enclosure containing the correlators were connected temporarily to two outputs of one compressive receiver. Using a 60 MHz CW input as a test signal, the correlator phase compensation line length (~ 45^{o}) was adjusted to produce a null at the quadrature correlator output.



(a)



(b)

Fig.7-4: Two views of the SAW coherent processor. In (a) the processor (top unit) is shown with its associated power supply. An internal view of the instrument is shown in (b). The covers are removed from the detector (left foreground) and correlator (right) assemblies. On the rear panel the LO chirp generator is on the right and the SAW device enclosure is visible at the lower left. The remaining processor modules are obscured by the sub-panel holding the detectors and correlators.

A check was made to ensure that the correlator output remained nulled across the spectrum analyser passband. The detector phase compensation leads were cut to the same length as the correlator line and a final check (still using one receiver output) made to verify that each detector output and the in-phase correlation product tracked across the passband.

With the correlators and detectors aligned and re-connected to two compressive receivers, the effect of the phase offset between the SAW convolvers was evident. Partial compensation of this and other phase errors (perhaps due to LO and output splitting network phase imbalance as well as small differences in cable lengths) was achieved by inserting a differential length of coaxial cable in the lines carrying the LO chirp to the signal mixers. The phase tracking of the entire processor was then checked and found to be satisfactory. Phase compensation of wideband RF systems is discussed further in section 10.4.

7.6 LABORATORY TEST RESULTS

Power responses of the two compressive receivers are shown in figure 7-5. The difference response between the channels is also plotted on the same axes. The curves were produced by applying a levelled sweep generator to the receiver inputs and using the digital spectrum integrator described in chapter eight to digitize and integrate the square-law detector outputs. If the sweep generator is swept asynchronously relative to the timing cycle of the SAW analyser, the integrated output reflects the compressive receiver passband shape. In effect, a pseudo-noise spectrum is produced, the advantage over a true noise spectrum being that the signal source is levelled accurately.

The -3 dB bandwidth of the compressive receivers is close to 30 MHz and the response characteristics of the convolvers allow a duty cycle of 60% (4.0 μs impulse PRI) before the output window becomes corrupted by aliased components (section 7.3). At 60% duty cycle, aliased signals are at least 15 dB below the main response in resolution cells at the edges of a 32 MHz output window (the width digitized by the prototype spectrum integrator) and are insignificant in other cells.

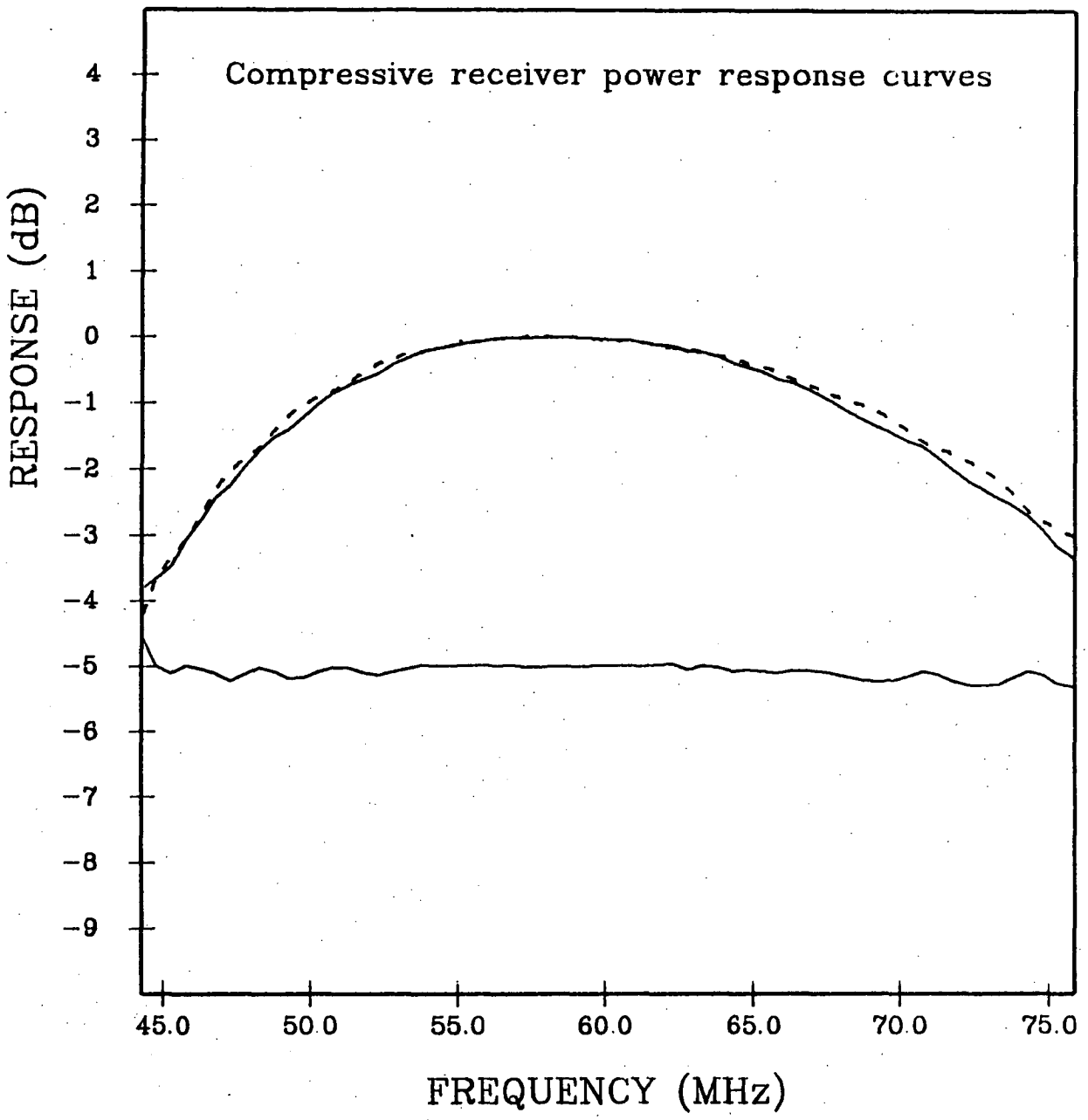


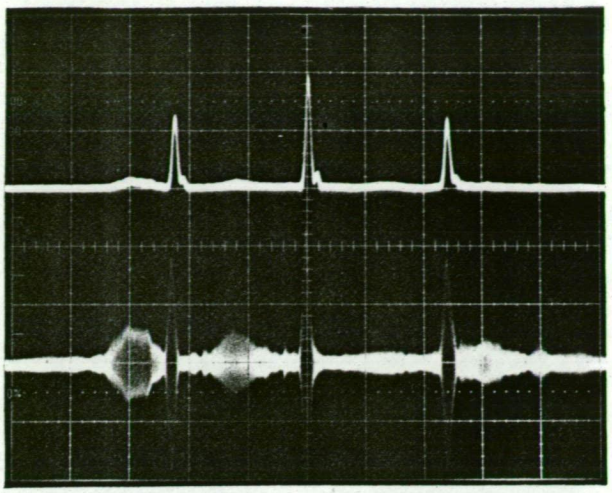
Fig.7-5: Power responses of the compressive receivers. The differential power response (offset by 5.0 dB) is shown below the main curves.

The taper of the passband is not typical of SAW analysers and is due to the low-cost DDL's used in the prototype instrument. The differential power response between the two receivers does not exceed 0.5 dB over the half-power bandwidth. Although not shown in figure 7-5, the differential response is within 1 dB across the input frequency range producing any significant output from the spectrum analysers. Coupling between the two channels is no greater than -40 dB over the entire operational range.

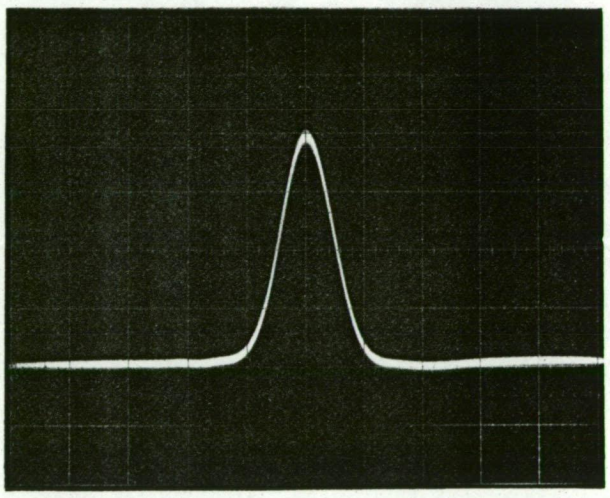
Pre-detection compressed pulses from one compressive receiver are shown in the lower trace of figure 7-6(a) and the corresponding square-law detected outputs (using the basic multiplier of figure 6-10) are shown in the upper trace. The adverse effect of the ripple compensated convolver on sidelobe performance is evident. Note the slight overshoot on the detected output due to the oscillatory impulse response of the third-order Butterworth post-detection filter. Figure 7-6(b) shows the detected output using the direct-coupled multiplier (figure 6-11). A Gaussian post-detection filter is used in this module, giving an improved time response.

The pre-detection single tone dynamic range characteristic of the compressive receivers is plotted in figure 7-7 and is of the expected form: at very high levels the output is clipped and at very low levels the output signal becomes submerged beneath the noise floor. A usable dynamic range of more than 40 dB is indicated. Two-tone, multi-tone and noise tests indicate a dynamic range specification of 30 dB to be a more realistic figure, the reduction being due to intermodulation in the pre-convolution signal stages and in the SAW convolvers.

Detection linearity using the basic detector (figure 6-10) is shown in figure 7-8. The effective dynamic range is about 25 dB, being governed largely by the 1.5 mV of video noise present at the output of the detector module. Over a 25 dB range the post-detection output remains within 5% of the ideal square-law characteristic. When the direct-coupled detector (figure 6-11) is used the dynamic range is restricted deliberately to 13 dB. The linearity over this range is excellent, being essentially the low-signal end of the characteristic given in figure 7-8.



(a)



(b)

Fig.7-6: Compressive receiver outputs. In (a) pre-detection (lower) and post-detection (upper) compressed pulses are shown. The three CW input signals are levelled to within about ± 1 dB and the basic correlator module of figure 6-10 is used as the square-law detector. Horizontal scale: 500 μs/div. Vertical scales: 0.2 V/div. (upper), 0.5 V/div. (lower). In (b) the post-detection output from the first video stage of the direct-coupled correlator module of figure 6-11 is shown. The input frequency is 60 MHz. Note that the Gaussian filter in the direct-coupled module gives a significantly better time response than the third-order Butterworth type used in the basic module. Horizontal scale: 50 ns/div. Vertical scale: 50 mV/div.

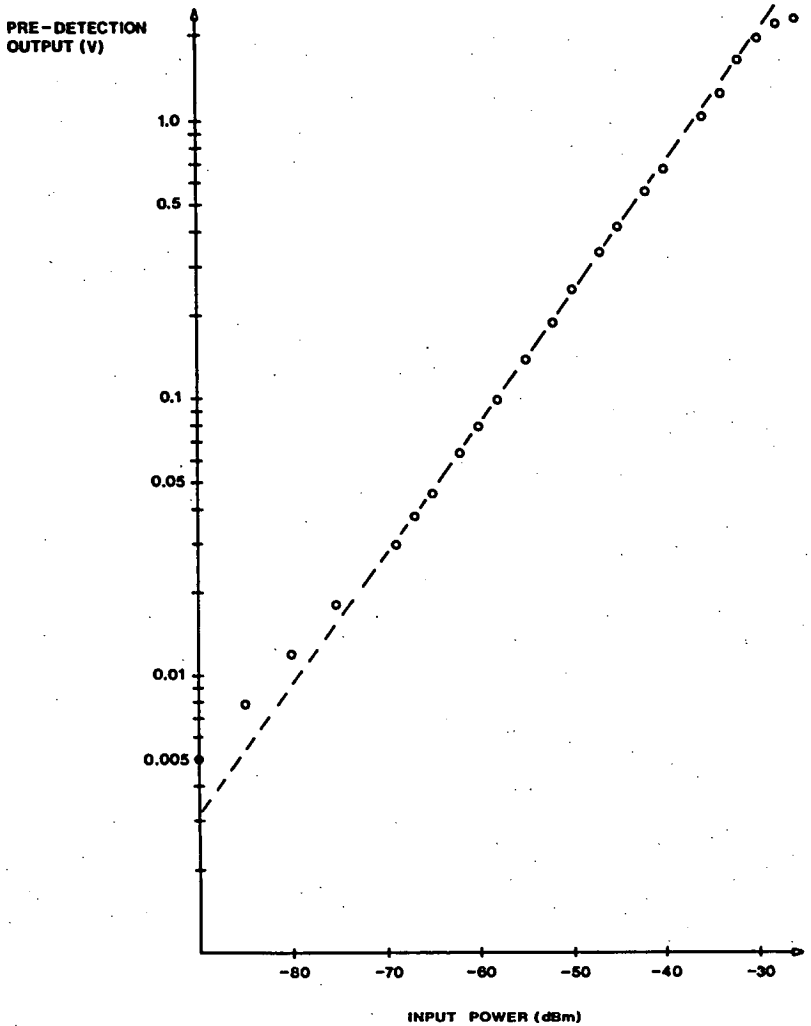


Fig.7-7: Compressive receiver pre-detection output voltage (measured at the monitor port of the signal distribution network) plotted as a function of input power for a 60 MHz CW test signal.

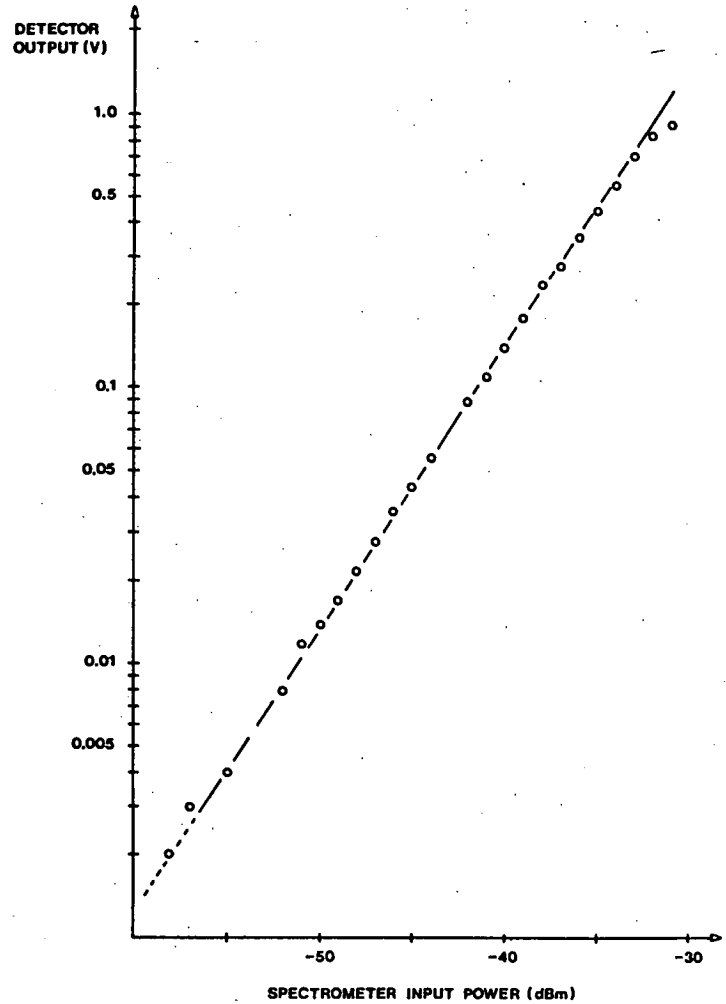


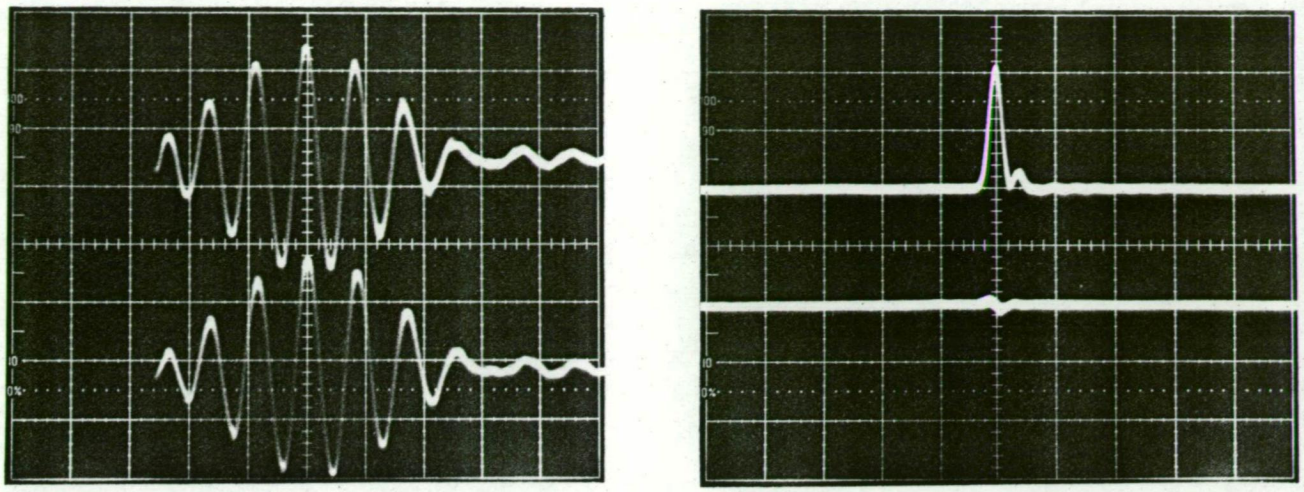
Fig.7-8: Compressive receiver post-detection output voltage versus input power using the basic correlator module shown in figure 6-10 as a square-law detector (60 MHz CW input). The slope of the line is consistent with a detection law of 2.05.

The potential of a SAW processor in applications requiring coherent signal processing is illustrated in figure 7-9. The photographs show the actual RF waveform of compressed pulses representing CW inputs to the instrument. The resulting outputs from the in-phase and quadrature correlators are also shown. In figure 7-9(a) the two input signals are in phase, leading to in-phase outputs from the two compressive receivers. In figure 7-9(b) the input phase difference is 180^{o}, and the RF components of the compressed pulses are now in anti-phase. As expected, the in-phase correlator output changes sign while the quadrature channel output remains zero.

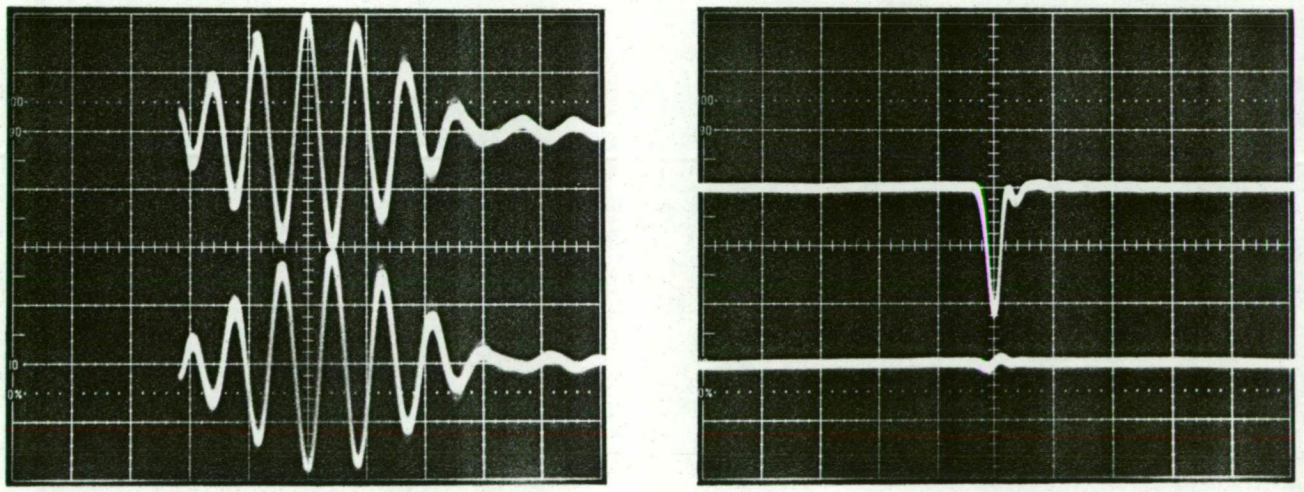
The form of the quadrature channel null is predictable from intuitive reasoning. Correlating two compressive receiver outputs is similar to the common process of correlating band-limited white noise in more conventional astronomical instrumentation. References such as Rogers (1976) give the shape of the in-phase and quadrature correlation products in conventional systems. The tapered compressed pulse spectra produce correlator outputs of similar form to those resulting from band-limited noise but without the leading and lagging ripples characteristic of rectangular spectra.

The correlation accuracy of the prototype IF processor is illustrated in figure 7-10 where the output voltage of both correlators is plotted as a function of input phase difference. The plots were obtained using an octave-band variable phase shifter in one IF input line. A continuously-variable fractional decibel attenuator was used in the other line to compensate for the insertion loss of the phase shifter. The amplitude and phase difference between the IF inputs was monitored using an HP 8405A vector voltmeter ($\pm 1.5^{o}$ phase accuracy), the sampling probes of which were connected directly to the processor front-panel inputs. Throughout the phase response measurements the attenuator was adjusted to maintain equal amplitudes at the two inputs.

Figure 7-11 shows the wideband correlation capability of the SAW instrument. With in-phase CW signals applied to the processor the in-phase product is indistinguishable from the intensity outputs and the quadrature product remains almost nulled. A high level of correlation accuracy is retained over a frequency range exceeding the



(a)



(b)

Fig.7-9: Coherent processor outputs. The left-hand oscillograms show pre-correlation outputs of the compressive receivers while the right-hand records show corresponding in-phase (upper) and quadrature (lower) correlator outputs. In (a) the outputs for an in-phase 60 MHz CW input signal are shown; in (b) the phase of one processor input has been shifted 180°. The horizontal scale in the left-hand oscillograms is 20 ns/div. The right-hand scale is 200 ns/div. Post multiplication filter: third-order Butterworth.

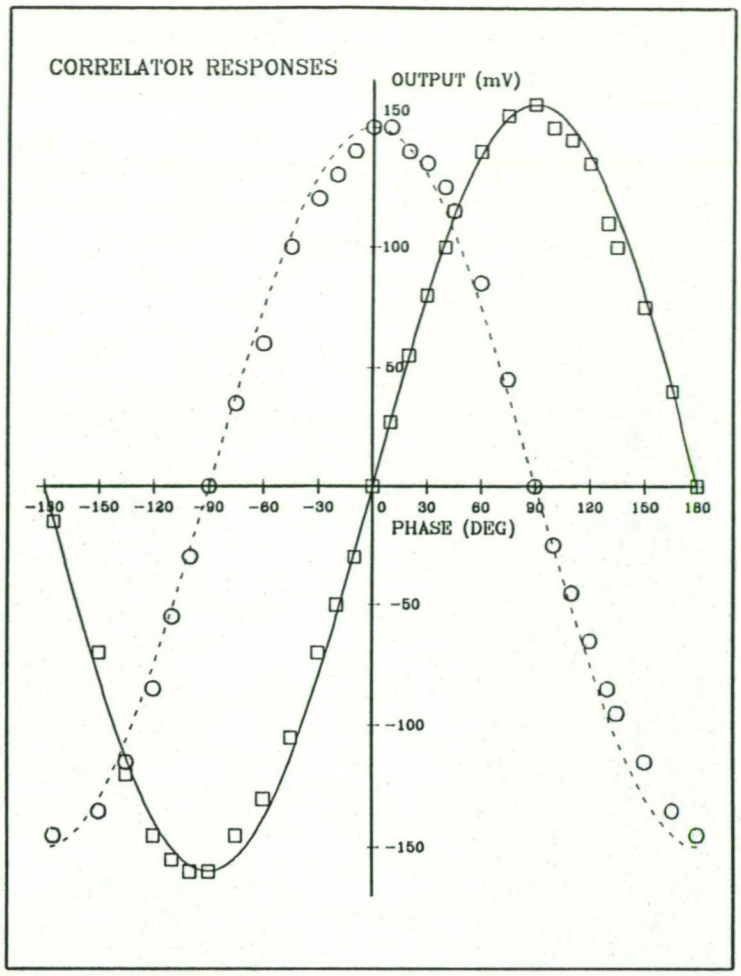


Fig.7-10: Phase response of the processor in-phase and quadrature correlation channels (60 MHz CW input). Ideal sinusoidal and cosinusoidal responses are also drawn.

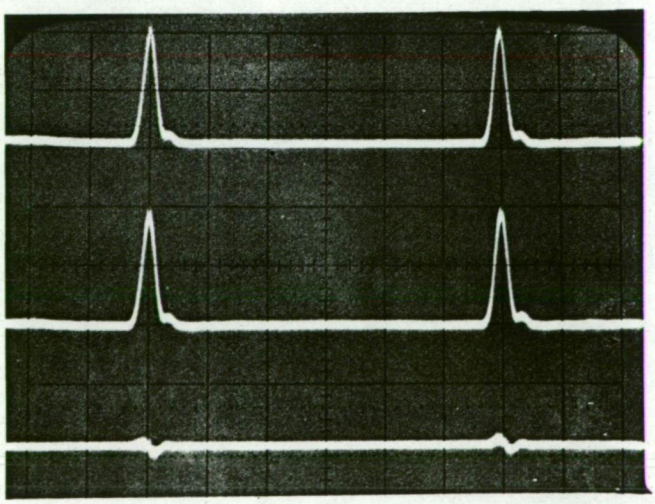


Fig.7-11: Processor output with in-phase 52.5 MHz and 67.5 MHz CW input signals. The upper trace shows an intensity output, the middle trace the in-phase product and the lower trace the quadrature product. Horizontal scale: 200 ns/div. Post-multiplication filter: third-order Butterworth.

nominal bandwidth. Figure 7-12 is a plot showing the residual phase error characteristic in more detail.

The self-squaring (self-detection) performance of the MM109 multiplier IC is illustrated in figure 7-13. The same CW signal was applied to both processor IF inputs. The top trace is the normal output of one square-law detector, the lower trace is the worst-case self-detection output of the second detector. Note that the unwanted output is about 1% of the wanted product. This figure does not change over the instrument passband. Self-detection occurs within the MM109 IC so both the basic and direct-coupled multiplier modules (figures 6-10 and 6-11) exhibit identical performance in this respect.

A value of 12.50 MHz/μs has been deduced for the output frequency-time relationship of the compressive receivers and, to a measurement resolution of 2 ns (0.04 resolution cell), no departure from linearity can be detected over the range 45-75 MHz. The group delays of the four processor output channels are well matched, the difference in the time of emergence of a variable-frequency test signal being less than the 2 ns resolution limit. This figure applies to both multiplier modules.

Chirp transformation is a linear process and the prototype compressive receivers should behave as linear devices over a wide range of input levels (figure 7-7). Nevertheless, a novel procedure was devised to verify specifically that the analysers exhibited linearity with CW and broadband input signals at both high and low signal-to-noise ratios. In addition, the test validated (at least qualitatively) the concept of an analyser resolution cell as an analog of a narrowband receiver. In the following description the term "signal-to-noise ratio" (SNR) is used in the communications engineering sense of mean signal plus noise to mean noise power rather than in the astronomical sense of mean signal to r.m.s. noise.

The linearity test was designed to allow direct comparison of the SNR in a selectable portion of a SAW analyser output with the SNR at the square-law detector output of a conventional receiver. Figure 7-14 shows the experimental apparatus used. The broadband

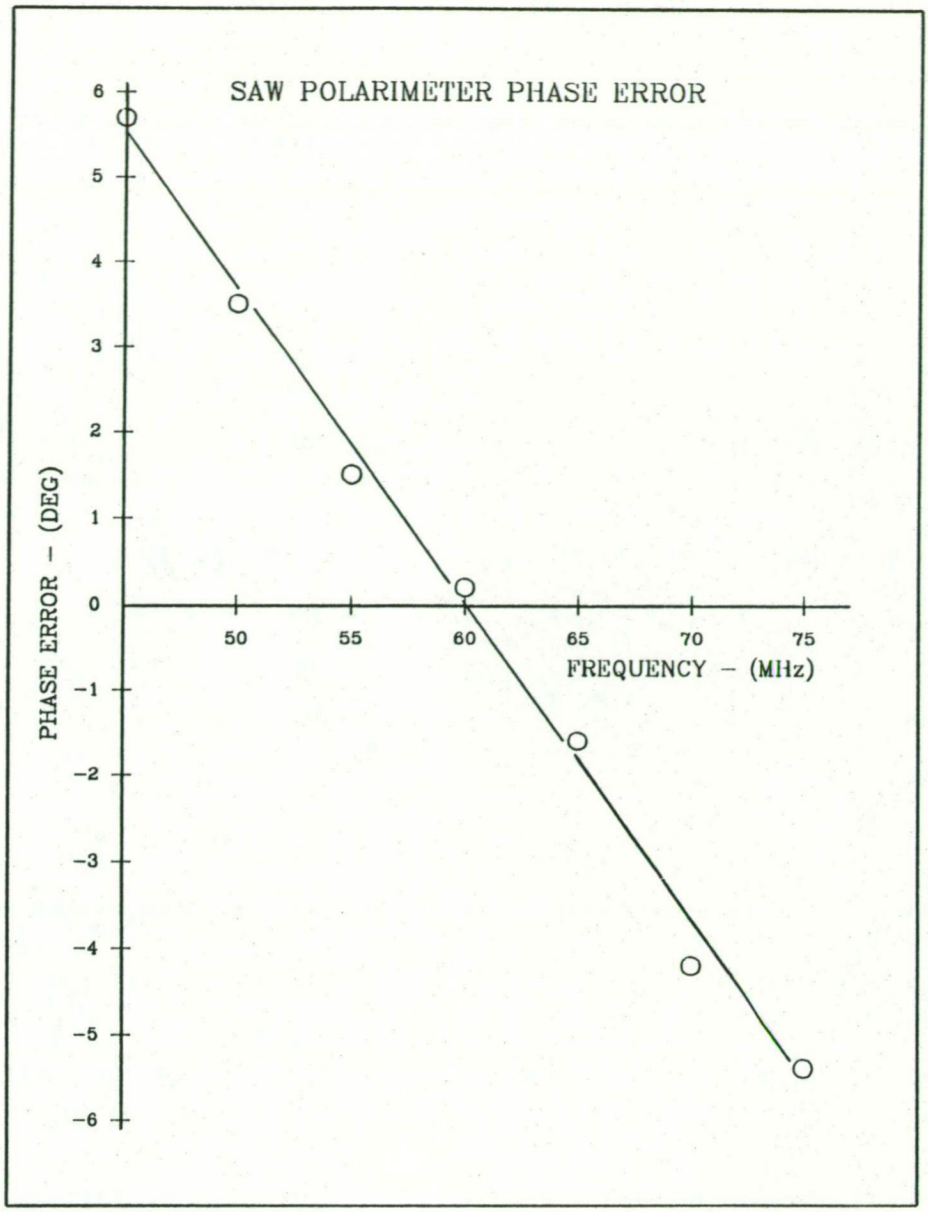


Fig.7-12: Processor phase error plotted as a function of frequency. The error bars at each point are ±1.5°.

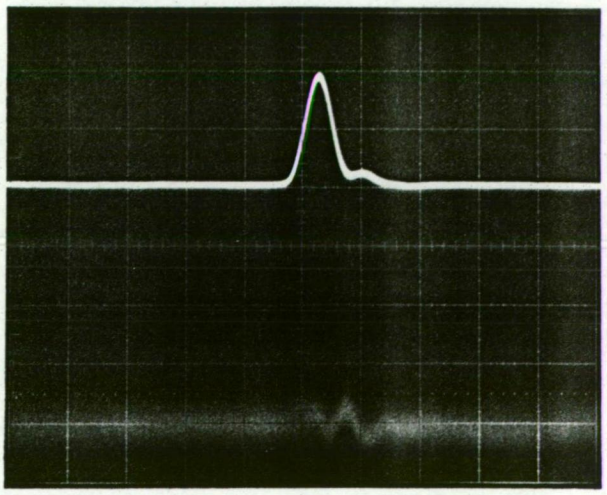


Fig.7-13: Self-detection characteristics of a multiplier module using the MM109 IC. The upper trace is the normal intensity output from one processor channel; the lower trace shows the worst-case self-detection output from the second channel. Horizontal scale: 100 ns/div. Vertical scales: 100 mV/div. (upper), 5 mV/div. (lower). Post-multiplication filter: third-order Butterworth.

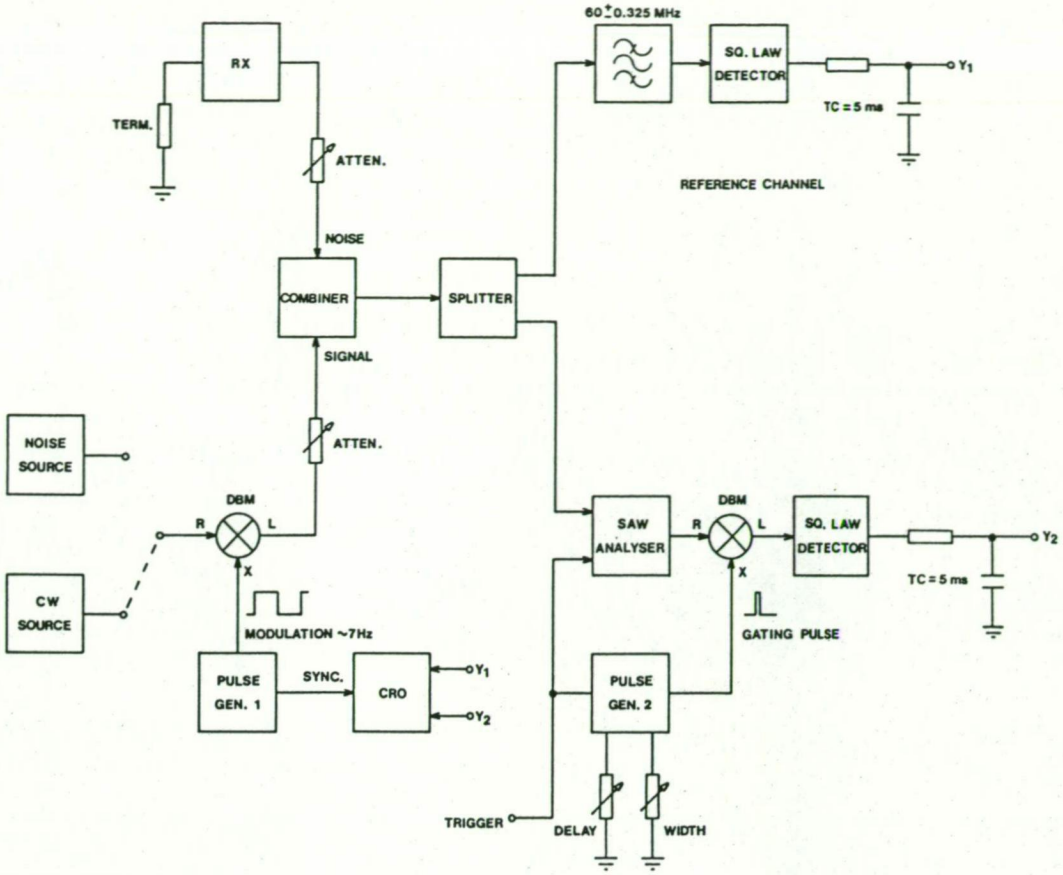
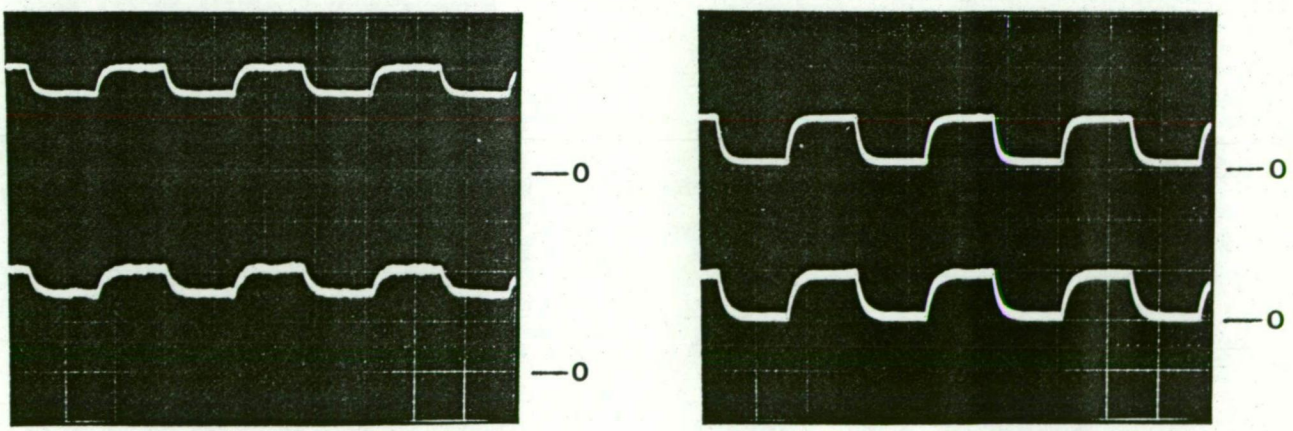


Fig.7-14: Apparatus for testing processor linearity.



(a) (b)

Fig.7-15: Examples of the output from the test arrangement shown in figure 7-14. In both (a) and (b) the upper trace shows the reference channel output and the lower trace the output from the gated SAW system (t_W~ 50 ns). The signal shown in (a) is a CW type while that in (b) is broadband noise.

receiver described in section 10.4 was used as a source of background noise, the receiver attenuator providing an easy means of adjusting the noise power. The signal was either a CW or broadband noise type, both of which were keyed at a low frequency. A narrowband filter (60.0 MHz centre frequency) with a power response approximating that of the SAW analyser resolution cell (figure 7-6(b)) was constructed, so the bandpass characteristics of the SAW and conventional systems were similar.

During the experiment the SAW spectrometer was triggered repetitively at a 6 μs trigger interval, the trigger pulse also being used to trigger a pulse generator. The generator output was used as a gating pulse for a high performance mixer, allowing a portion of the spectrum analyser pre-detection waveform to be excised. The pulse generator "delay" and "width" controls enabled the position and extent of the excision gate to be adjusted. Great care was taken to ensure that the width and position of the excised window remained correct at all stages of the test.

The selected section of the spectrum analyser output was detected, then integrated in an RC integrator. For comparative purposes the integrator time constant was set equal to the standard channel post-detector time constant. This is not necessary for the linearity evaluation where the only requirement is that the outputs need to be smooth enough to allow unambiguous determination of mean signal and noise levels. However, the use of equal time constants allowed an approximate assessment of the relative SNR performance (in the astronomical sense) of the two systems.

The dual-trace CRO display of the integrated outputs was synchronized with the signal modulation source and, when photographic records were made, the display gains were equalized by making the range from no signal (maximum receiver and signal attenuation) to maximum output (that obtained with a moderate level, noise-free CW signal) occupy the same range on both displays.

With a CW input the SAW system and the reference channel produced the same output over all likely combinations of input signal and background noise levels, provided that the gating window width (t_w) did not exceed the resolution cell width of 50 ns. For $t_w \gtrsim 50$ ns the SNR relative to the standard channel decreased since the noise

contribution at the output was increased with no commensurate increase in signal power. With $t_w \lesssim 50$ ns the absolute output from the SAW system integrator decreased but the SNR remained constant. This demonstrates the validity of using the resolution cell width as a measure of bandwidth in SNR calculations. Figure 7-15(a) is an oscillogram of the two analog integrator outputs obtained with a fairly low-level CW input ($t_w \sim 50$ ns).

With a pulsed noise signal the SAW analyser was again linear at both high and low SNR. For $t_w \lesssim 50$ ns the output behaviour was as for the CW case but, since both the signal and background were broadband, the output SNR also remained constant for $t_w \gtrsim 50$ ns. Figure 7-15(b) shows the two outputs with a high SNR broadband pulse used as an input signal ($t_w \sim 50$ ns).

Although designed primarily to check linearity (relative signal and noise levels), the test can be used to gauge approximately the SNR (used from this point in the astronomical sense) performance of the two systems. In figure 7-15(a) for example, the SAW analyser is noisier than the reference channel by a factor of about 1.9. At 6 μs trigger rate the compressive receiver duty cycle is 40%, leading to an expected SNR loss factor of 1.6. Given the likely error in matching the frequency response of the standard filter to the shape of the compressive receiver resolution cell and the error in estimating r.m.s. fluctuations, the 19% difference is acceptable.

The linearity test is an important one and it is worth outlining practical aspects of the spectrometer gating process. The pulse generator and DBM must be high quality types, HP 8013B and HP 10514A models being used to produce the results shown earlier. The gating pulse level and d.c. offset need careful adjustment for optimum switching speed and isolation characteristics. The DBM must be terminated properly at all ports. Even with optimum operation, pulse generator and DBM limitations combined to make the effective gate width in the real system rather uncertain for $t_w \lesssim 50$ ns.

7.7 CONCLUSION

The results of the coherent SAW processor tests are encouraging and the novel aspects of the design work well in practice. The poor sidelobe performance due to the use of reciprocal ripple compensated convolvers is inherent in even the simplest compressive receiver based on DS1301/1302 DDL's. With more suitable SAW devices the sidelobe performance would be improved markedly. In fact, uncompensated convolvers would be easier and cheaper to produce than the DS1302 devices.

The high sidelobe levels in the prototype processor imply a higher than usual degree of cross-correlation between noise components in separate spectrometer cells. The effects on all types of astronomical observation are difficult to quantify but in the case of pulsars for example, qualitative studies by Boriakoff (1973) and Hall (1980) indicate that noise correlation is detrimental to de-dispersed SNR when the pulsar dispersion smear time (Δt_d) is significantly less than $1/\Delta f$, where Δf is the frequency resolution cell width. With a resolution cell width of the order of 0.65 MHz in the prototype processor and an observing frequency of 650 MHz, a pulsar of DM ~ 0.1 would be needed to produce $\Delta t_d = 1/\Delta f$. Real pulsars have dispersion measures at least an order of magnitude greater than this figure so $\Delta t_d >> 1/\Delta f$ in all cases. Thus, so far as sensitivity is concerned, the instrument is affected only slightly by the presence of sidelobes. More generally, it is doubtful if the sidelobes are high enough to degrade significantly the effective frequency resolution during observations of dynamic radio emission, particularly narrowband emission such as solar bursts.

The phase tracking of the processor is considered satisfactory. More elaborate phase balancing methods could be employed but, with only $\pm 6^{o}$ linear phase error over 30 MHz, the SAW processor is a usable radio polarimeter. Calibration techniques undertaken during observations allow the removal of residual phase error as well as correction for non-uniform spectrometer amplitude response. It is worth noting that performance of the coherent processor in a radar

intercept or similar role would be outstanding since accurate, high time resolution angle and time of arrival information is obtainable from a receiver which is much simpler than conventional multi-beam arrangements.

CHAPTER 7 - REFERENCES

Boriakoff, V.: Ph.D. thesis. Cornell University, 1973.

Hall, P.J.: B.Sc.(Hons.) thesis. University of Tasmania, 1980.

Rogers, A.E.E.: Interferometers and Arrays - in "Methods of experimental physics". Vol.12, Ch.5, Academic Press, 1976.

CHAPTER 8 A DIGITAL VIDEO INTEGRATOR FOR SAW SPECTROMETERS

8.1 INTRODUCTION

If SAW spectrum analysers are to be useful in radioastronomy and other applications where computer acquisition of spectral data is a requirement, some form of intermediate signal processor must be interposed between the spectrometer and data acquisition computer to reduce the output data rate to one compatible with the transfer and processing capabilities of existing digital computers. Although A/D conversion and digital processing is not trivial at the high sample and clock frequencies necessary, the intermediate processor will probably be a digital instrument because the alternative of fast analog processing is barely feasible.

In section 4.8 two methods of data rate reduction using digital methods were examined and the conclusion reached that the appropriate type of processor in astronomical applications is a spectrum integrator. The function of an integrator is to add synchronously a set number of successive spectra from a SAW spectrum analyser. The detected or correlated analyser output is at baseband so the term "digital video integrator" (DVI) has been adopted to describe the intermediate processor.

Integrating the spectrometer output degrades the time resolution but reduces the standard deviation of spectral estimates in a given frequency resolution cell by a factor of $\sqrt{N}$, where N is the number of integrations. An alternative view is that the output bandwidth of the spectrometer-integrator combination is reduced by a factor of N relative to the spectrometer bandwidth alone.

A DVI is analogous to the post-detection integration networks in a filterbank spectrum analyser and in figure 8-1 the analogy is extended to illustrate the function of the digital instrument. As the time waveform (or spectrum) emerges from the SAW spectrometer, successive RC networks are connected sequentially to the detector output. The commutation rate and the number of integration stages is arbitrary but is usually such that there is at least one stage

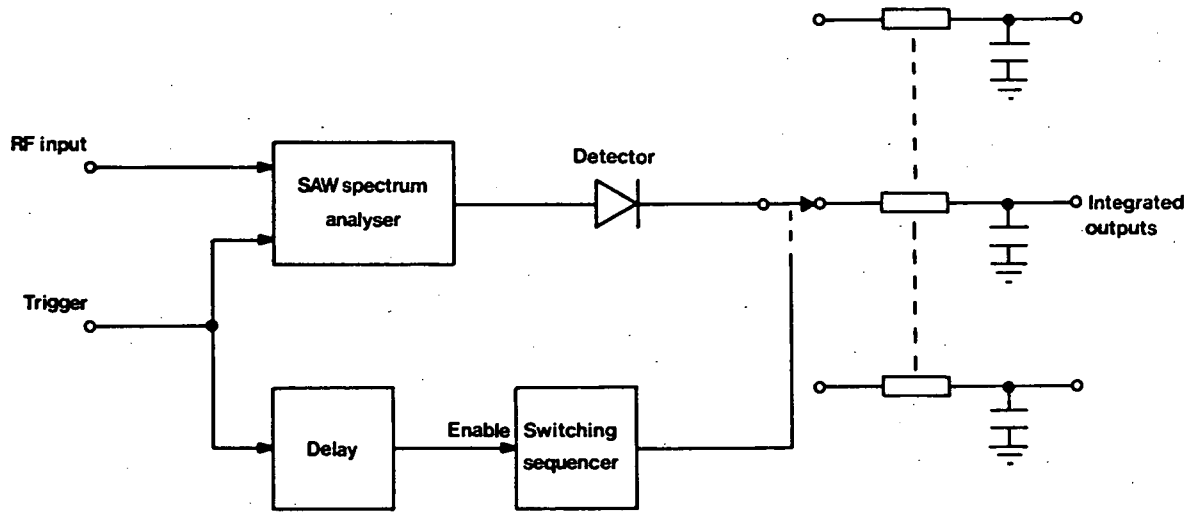


Fig.8-1: Analog video integrator.

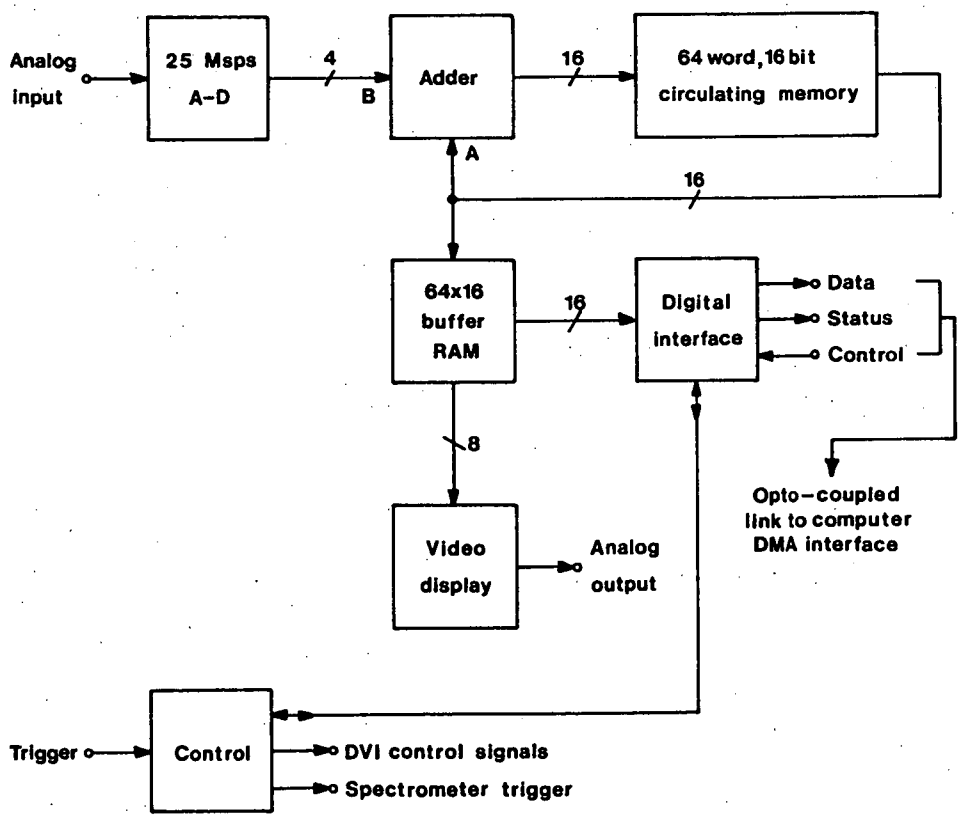


Fig.8-2: Simplified block diagram of the "mark 1" digital video integrator. The computer interface is described in section 10.3. The maximum sample rate is 30 mega-samples/second but 25 Msps operation is normally used (section 8.4).

per spectrometer frequency resolution cell (section 4.9). If this criterion is met exactly the commutation frequency is equal to the spectrum analyser output bandwidth. The spectrometer-integrator combination (SAW radiometer) therefore resembles closely a standard filterbank analyser. In a polarimeter or similar SAW coherent processor one spectrum integrator is required for each of the four processor video outputs.

This chapter describes a prototype DVI designed to integrate one output of the coherent SAW processor described in chapter seven. Four DVI's have been constructed, allowing all four processor outputs to be integrated simultaneously. The DVI's share common control circuitry and are contained in a single rack-mounted enclosure. Digital representations of the four integrated spectra are time division multiplexed, then transmitted to a data acquisition computer via the high speed interface and data link described in chapter ten. Analog representations of integrated spectra are always available and, with a suitable CRT display, the four spectra can be displayed simultaneously regardless of whether a computer is used to acquire spectral data from the integrators.

The DVI (referred to in the text as the "mark 1" instrument) has the capability of sampling and processing at 30 MHz, three times as fast as the faster of two earlier spectrum integrators (Alldritt et al., 1978; Racal-MESL, 1982). In contrast with the other units the DVI is compact and cheap. The present system is the first to provide the option of computerized data acquisition and is apparently the first to use a simple (yet satisfactory) triggering system to control the SAW analyser and integrator. The instrument provides continuous spectral integration and no observing time is lost during data transfer or display operations. This last feature is the main one distinguishing the DVI from very expensive, fast waveform analysers available commercially. Many analysers are simply transient recorders while more elaborate instruments offering signal processing facilities usually require stored data to be down-loaded for operations such as integration.

8.2 CHOICE OF INTEGRATOR TECHNOLOGY

In principle the functional diagram of figure 8-1 could serve as a basis for an analog spectrum integrator. However, the difficulties in realising such an instrument are great. The commutation rate is high (20 MHz for the prototype SAW analyser if one sample per resolution cell is taken) and the practical limitations of analog technology make it difficult to transfer small quantities of charge without introducing switching noise.

After consideration of the limitations of spectrum integrators using analog or CRT-based electro-optical methods, digital technology was adopted as the basis of practical instruments. Although high A/D conversion rates are necessary, once data are encoded into digital form there is essentially no data deterioration in subsequent manipulation and transfer operations.

In both the "mark 1" DVI described in this chapter and the larger memory capacity "mark 2" instrument mentioned in appendix five, extensive use has been made of the new Fairchild "FAST" or "F" series TTL logic family. With flip-flops and counters able to be clocked at 150 MHz and gate delays of 2-4 ns, "F" series logic systems can approach ECL systems in terms of speed. In addition, "F" series TTL (or its equivalent in other manufacturers' product lines) offers the following advantages over ECL:

(a) a factor of four (typically) increase in noise margin;
(b) lower power consumption (comparable with "LS" series TTL);
(c) pin-for-pin compatibility with, and direct interface to, standard TTL devices;
(d) large drive capability (comparable with "S" series TTL).

8.3 BASIC DVI OPERATION

The "mark 1" and "mark 2" integrators are quite different in internal operation but are functionally interchangeable in terms of the input and output signals described in this section. Both conform to design guidelines developed for astronomical instruments. The

guidelines recognize that SAW spectrum analysers are fast and digital computers are slow. In a real-time observing situation it is therefore desirable that the processing cycle of the high speed device should remain unaffected by operations of the slower unit.

The spectrometer-DVI combination is a stand-alone analysis system and does not require an external processor to be present for spectral analysis and integration to proceed. A computer may acquire integrated spectral data from the DVI upon receipt of permission from the DVI to do so. The computer may start or stop the spectrum analysis process but cannot interfere in the spectrometer-DVI operational cycle once analysis has been initiated.

Figure 8-2 is a simplified block diagram of the "mark 1" DVI and figure 8-3 shows the timing relationship between major signals. The DVI internal control signals and the spectrometer trigger signal are derived from a master crystal oscillator running at the A/D converter sample frequency. The control section has two trigger modes: internal and external. Selection of internal trigger means that the SAW spectrometer is triggered by a divided-down version of the master oscillator. In the "mark 1" instrument a 25 MHz master oscillator is divided by 100 to produce a 4 μs trigger period, the period necessary to maintain 60% duty cycle operation of the SAW processor described in chapter seven. Both the "mark 1" and "mark 2" DVI's are capable of supporting 100% duty cycle operation with no hardware changes.

External trigger mode forces the DVI to trigger the spectrometer at a rate determined by the period of a pulse train applied to the external trigger input. Alternatively, a single trigger pulse can be applied to digitize one data window. An external trigger facility is useful for test purposes and its inclusion allows the DVI to be used as a general purpose transient analyser as well as a dedicated instrument.

In external trigger mode the DVI synchronizes a trigger signal to the internal master clock, so the instant of spectrometer triggering could be up to one clock period later than the requested time. This locking together of the spectrometer and integrator timing is central to the design philosophy and ensures that the timing of the two instruments remains in step in all operating modes. If a particular

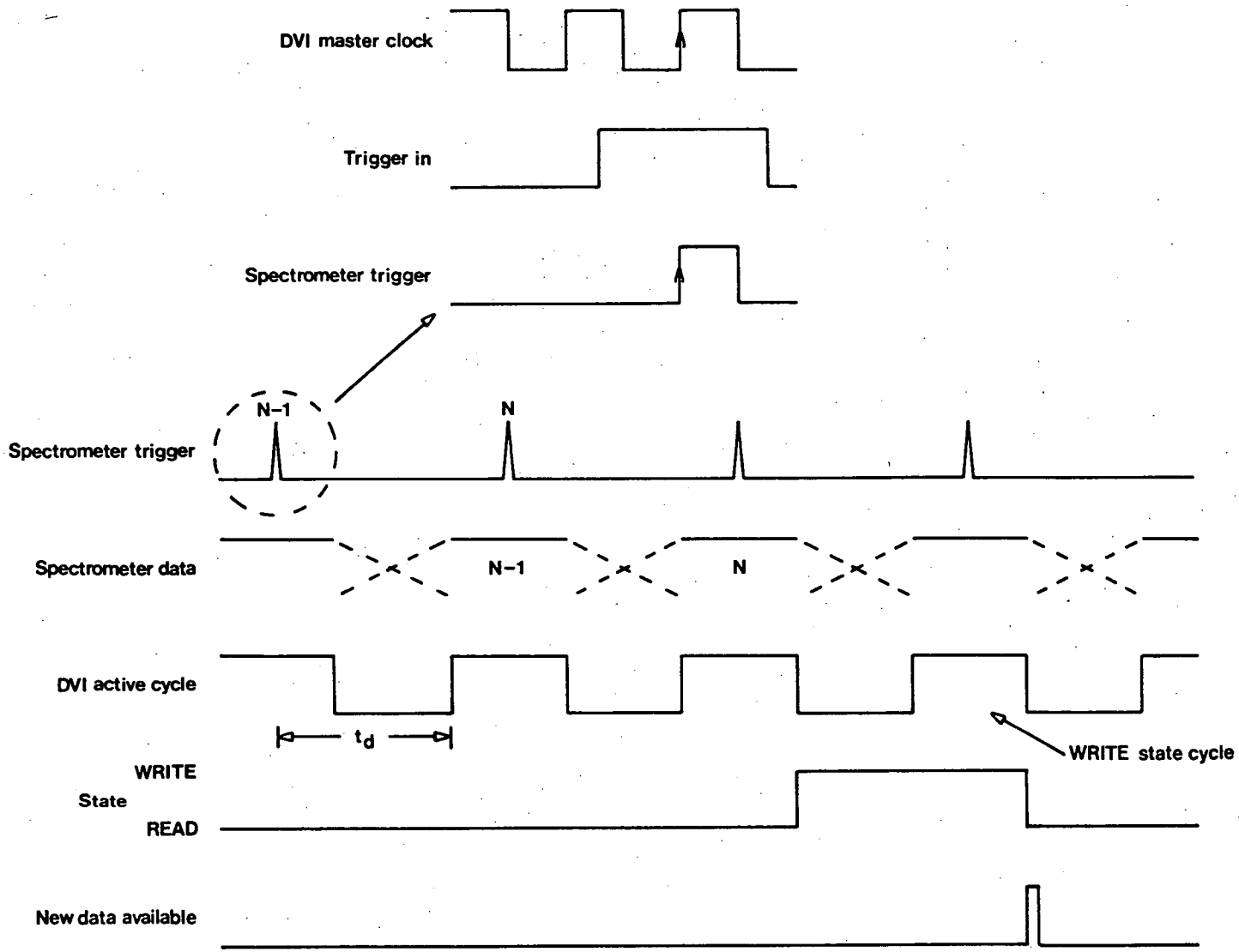


Fig.8-3: Basic DVI timing.

application calls for extensive use of the external trigger facility and the small amount of jitter (40 ns maximum with 25 MHz sampling) introduced by re-synchronization is unacceptable, the trigger can be derived from the DVI clock or both the DVI clock and trigger signals can be derived from a common external frequency standard.

The time between triggering of the SAW analyser and emergence of output data (t_d in figure 8-3) is constant and is determined by the absolute delays of the dispersive delay lines used in the SAW instrument. The DVI control section contains an absolute delay generator which prevents integration beginning until the spectrometer full-compression data window appears.

When time t_d has elapsed, integration is initiated. An accumulated sum is read out of a circulating memory at a rate equal to the A/D converter sampling rate. The timing is arranged such that when data representing the integral in spectral cell x are present at the adder "A" input (figure 8-2), newly converted data from cell x are present at the "B" input. The same operation that reads the old sum for cell x+1 from the end of the memory also writes the newly formed cell x sum (old sum + A/D converter output) back into the beginning of memory. At the end of the sequence, data for the first spectral cell are at the end of the memory and the integration cycle can re-commence when initiated by the absolute delay generator.

The number of spectra to be added (N) is set on thumbwheel switches located on the DVI front panel. The maximum value of N is 9999 giving an integration time of about 20 ms with the "mark 1" DVI if continuous full-scale output from a four-bit A/D converter is assumed. Note that with the prototype SAW spectrum analyser operating at 60% duty cycle, the elapsed time for an observation is 1.67 times the effective spectral integration time. In a line spectrometer application, a DVI can provide all the required integration if the facility to set larger values of N is included. This modification is easy to incorporate into existing DVI designs. Additionally, extension of present 16-bit memory designs to the 32 or more bits likely to be necessary in such an application is simple.

After N integration cycles have been completed, the DVI enters a "write to buffer" or WRITE state. During this state the loop between the circulating memory and the adder is opened and, in the active part of the state (referred to as the "WRITE state cycle"), summed spectral data are fed to the output buffer RAM rather than to the adder. The adder "A" input is grounded, causing zero to be added to the A/D converter data. A/D data are therefore sent directly into store and are ready to be used as the old sum for the next integration cycle. Thus, the circulating memory is cleared automatically by the integration process. There is no additional time overhead involved in the WRITE state data transfer operation so integration is continuous.

At the end of a WRITE state, the DVI returns to its normal READ state where data in the output buffer are available for display and transfer to a data acquisition computer or other device. Upon entry to a READ state the integrator generates a "new data available" signal, indicating that new spectral data are ready and that an external device may read the output buffer. When the DVI enters the next WRITE state it immediately regains control of the buffer. A valid reading of the buffer by an external device occurs when the device acquires all data in the buffer before being interrupted again by the DVI "new data available" signal.

In normal operation the device reading the DVI output buffer is either the internal video display unit (which provides a CRO display of integrated spectral power versus frequency cell) or a data acquisition computer communicating with the DVI via a high speed interface. In the second case the CRO display is still available and, in addition to its primary function, indicates satisfactory data transfer since clock pulses necessary for the display operation are then supplied by the computer.

8.4 THE MARK 1 DVI

The "mark 1" integrator was the first digital processor to be constructed and served to validate the operational protocol developed for the SAW spectrometer, integrator and computer combination.

The major specifications of the instrument are listed in table 8-1.

Table 8-1 Mark 1 DVI Specifications

Sample frequency:	30 MHz maximum, 25 MHz typical.
Quantization resolution:	4 bits.
Number of frequency cells:	64.
Trigger modes:	Internal or external.
Operating modes:	Normal (continuous) or pulsar (start-stop).
Video input range:	-1 to 0 V.
Buffer readout options:	Local (analog) or line (computer).
Analog output range:	± 5 V.
Analog display facilities:	Digital gain control, switchable d.c. offset giving continuously variable offset when selected.
Buffer output logic:	Tri-state TTL, enabled externally to allow connection to data bus.
Power requirements:	+5 V (15 A), ±15 V (1 A), ±10 V (1 A) for four integrators.

Although the maximum sample frequency is 30 MHz, 25 MHz sampling is used with the SAW spectrometer described in chapter seven. This gives integrated spectral estimates in 64 frequency cells extending over a 32 MHz bandwidth. Hence, the SAW analyser output is digitized over an output window slightly wider than its -3 dB passband (figure 7-5). Adjacent samples are marginally more correlated than samples taken at 20 MHz, the nominal output bandwidth. The form of the spectral sampling function is discussed in section 9.2 and illustrated in figure 9-6. Pulsar mode operation is described in section 10.2 and is not considered in detail in this chapter.

The DVI has four distinct modules. These are the A/D converter, arithmetic and logic unit (ALU), output buffer RAM, and control signal generator. Each unit is physically self-contained in the final version of the integrator and the general arrangement of the basic instrument is as shown in figure 8-2. The architecture makes it simple to interleave integrators if very high effective

sampling rates are required. In practice, the circulating memory is formed from dual, 64 stage shift register IC's and the 64 word output buffer RAM may be either a fast, low density type constructed using sixteen 16 × 4 ECL-TTL RAM chips (section 8.7) or a slightly slower version using two higher density bipolar RAM IC's (Signetics N82S09). In the second case the maximum DVI sample frequency is reduced to 18 MHz but it is feasible to construct the entire instrument on a single PCB. Figure 8-4 shows a single board version of the "mark 1" DVI (actually the first prototype integrator).

The advantages of the final modular style of construction are that PCB changes to a given module do not require re-construction of the entire DVI and importantly, several A/D converter, ALU and buffer boards can be slaved conveniently to a common control board. Hence, all four DVI's required in a coherent SAW processing system need only a single control module and can be contained conveniently in a common enclosure.

The internal layout of the DVI enclosure is shown in figure 8-5. A rack-mounting box (430 mm × 265 mm × 400 mm) contains four A/D converters, four ALU and buffer modules, a single control module, a ± 10 V A/D converter power supply and the "remote" end of a computer interface and data link. Integrator PCB's (315 mm × 180 mm) plug into edge connector slots and control signal back-plane wiring radiates from a central slot in an arrangement which minimises timing skew between the four DVI's. The back-plane wiring also carries a 16-bit output data bus and provides low inductance power and ground distribution. The DVI +5 V and ±15 V power supplies are located in a separate rack-mounting box which also contains the SAW processor supplies. These power rails enter the DVI enclosure via a rear-mounted multi-pin connector. The mains lead for the A/D converter power supply plugs into a switched 3-pin socket on the back panel of the main power supply box, preventing the analog section of the A/D converters being powered alone. Forced air cooling of the DVI enclosure is provided since the heat generated by the high speed buffer RAM is considerable. Figure 8-6 shows the DVI front panel layout. The function of each control and connector is summarized in appendix four.

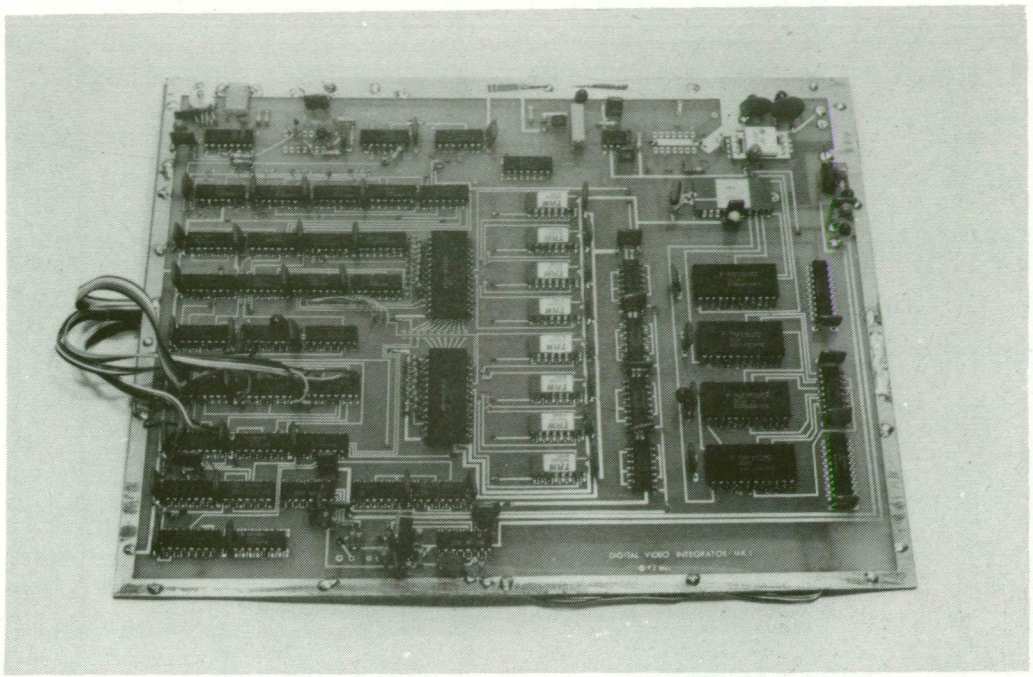


Fig.8-4: Single PCB version of the "mark 1" DVI.

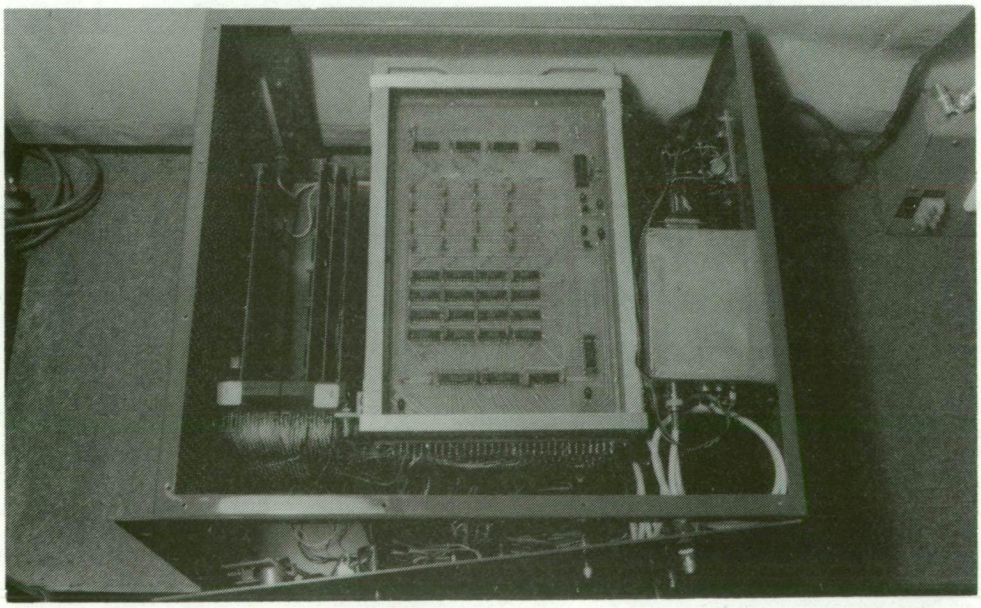


Fig.8-5: Internal layout of the DVI enclosure. The computer interface PCB and the data link circuitry is on the left. The central PCB mounting assembly contains four ALU and buffer boards and a single control board. The four A/D converter modules are located in the metal boxes on the right and the A/D converter analog power supply is at the right rear.

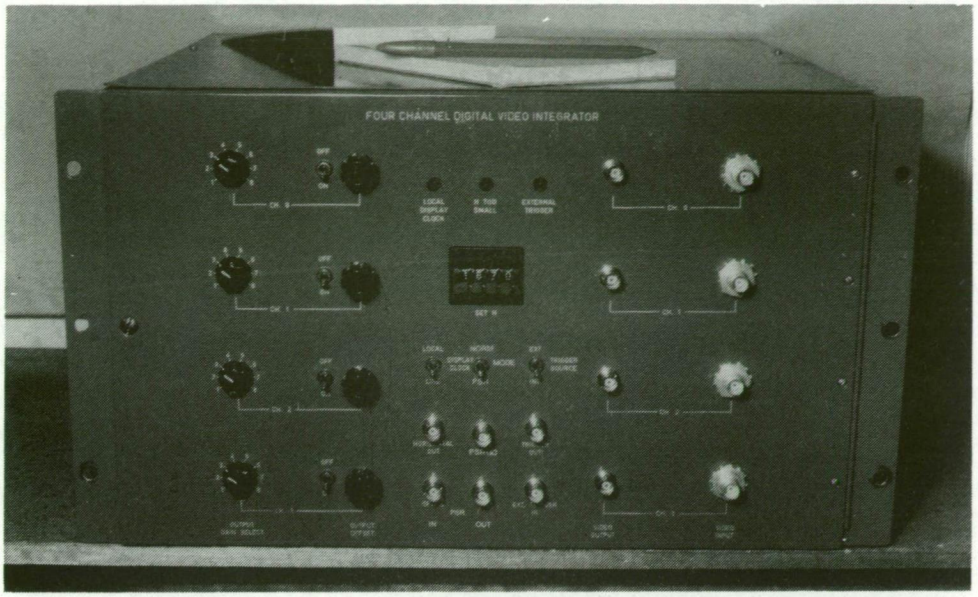


Fig.8-6: DVI control panel. See appendix four for a description of the function of each control and connector.

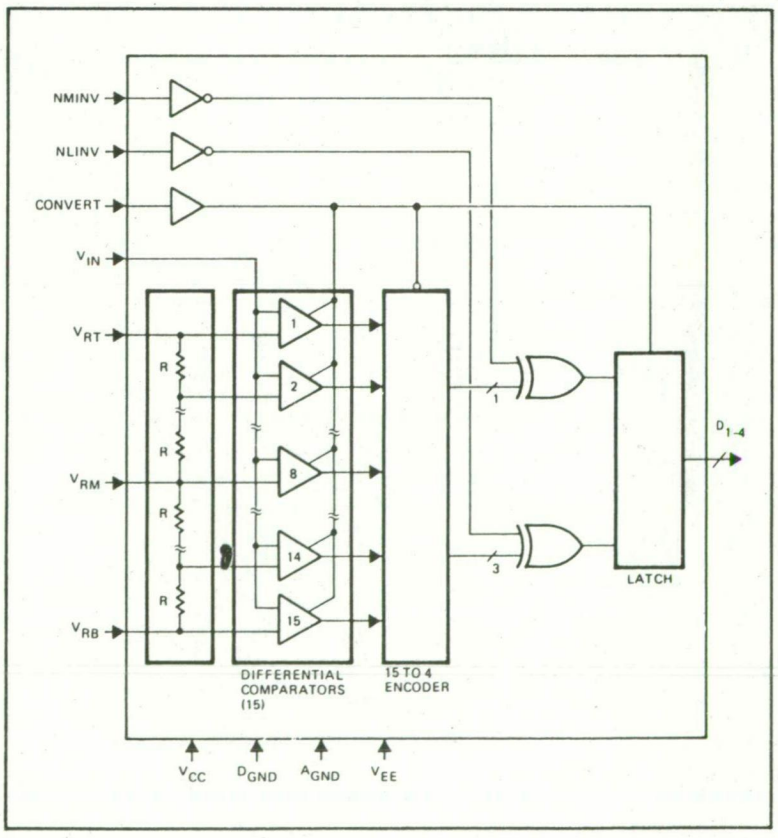


Fig.8-7: Block diagram of a typical flash A/D converter. The diagram is of the TRW TDC1021J four-bit device used in the "mark 1" DVI.

Although the DVI operation is complex in detail, a brief description of each of the four modules is now given. Where appropriate, major design considerations for various sub-sections are discussed. The descriptions are included because some techniques used in the DVI design and construction are unusual, being more appropriate to high speed logic systems than to conventional systems. However, readers wishing to avoid the engineering descriptions can proceed to section 8.9 with no loss of textual continuity.

8.5 ANALOG - DIGITAL CONVERTER

The most important requirement of the A/D converter is that it be capable of converting at the high rate necessary in a DVI. For a sample frequency of 30 MHz, the absolute maximum conversion time is 33 ns. Successive approximation A/D converters are too slow for use in high speed video applications and, although "propagation" type converters can be made to function at up to 20 MHz clock frequency (Zuch, 1979), modern bipolar VLSI technology allows implementation of "flash" converters suitable for use where conversion speeds of several tens of megahertz are needed (Swartzlander, 1982). Use of the flash technique in conjunction with hybrid ECL or GaAs technology makes it possible to produce three or four-bit A/D converters capable of sampling at 500 MHz or more. The simulations described in chapter five indicate that a four-bit converter is satisfactory for most applications in astronomy.

Figure 8-7 is a block diagram of a flash A/D converter. The converter consists of a quantizer followed by an encoder. Most converters also have an internal latch to staticize encoded data. The quantizer is constructed using a bank of 2^{n-1} analog comparators, where n is the number of output bits. Each comparator compares the analog input signal with a divided version of a reference voltage. Depending on the input voltage, all the comparators up to a certain level are on, while those above that point in the comparator chain are off. The transition point is encoded in the output encoder, a purely combinatorial logic circuit designed to convert the Johnson code at the comparator outputs into a more readily usable code such as straight binary or two's complement.

Regardless of the method of A/D conversion, there are several performance-limiting parameters to be considered when choosing a video converter. These parameters are in addition to the usual d.c. and low frequency considerations such as linearity, temperature sensitivity, offset errors and power supply sensitivity. Most of the high frequency errors are aperture related in the sense that they are linked to the non-zero width of, or the variation in the width or position of, the sampling window or aperture. In the past it has been necessary to use high speed sample-and-hold circuits ahead of video A/D converters to minimise high frequency errors. Many flash A/D converters are now fast enough to allow the designer to dispense with the sample-and-hold, even in wideband video systems.

The first high frequency characteristic of interest is the sampling time offset. It is the time taken for the A/D converter to begin a valid conversion and is measured from the active edge of the "convert" signal to the start of the conversion process. The offset does not in itself cause errors in the digitization process since it can be compensated for by changing the phase of the A/D converter clock relative to the phase of other signals in the digital processor.

Aperture time is a second high frequency parameter and is the time taken for the A/D converter to examine the analog input in order to make a conversion. Non-zero aperture time produces errors in digitization if the analog signal changes during the aperture window. The maximum error is related to the highest rate of change of the input signal and, assuming a full-scale sinusoidal input, a well known result (see, for example, Zuch 1979) can be modified to give the relationship between aperture time and error as:

$$t_A = \frac{2^{x-1}}{2^n \pi f} \qquad (x > 0) \qquad 8.5.1$$

where t_A is the aperture time, n is the number of bits encoded, x is the error in bits and f is the input frequency. For a 20 MHz input signal the aperture time of a four-bit A/D converter must be less than 1.0 ns if the aperture error is not to exceed one bit.

Aperture jitter is a third important high frequency parameter and is defined as the short-term standard deviation of the sampling

time offset. A related process is the cycle-to-cycle jitter of the sampling clock. Bird and Folchi (1982) show that there is a direct correlation between the magnitude of jitter processes and amplitude errors in A/D conversion. The error mechanism can be visualized by considering successive samples at the same nominal phase on a periodic waveform. Jitter causes each sample to be slightly different and the magnitude of the error is clearly related to the rate at which the input signal can change over the jitter time-scale. To quantize a 20 MHz sine wave with a peak inaccuracy of 1/2 LSB, the sampling time-base jitter must not exceed 0.5 ns (peak) for a four-bit A/D converter.

Sampling jitter limits can be used to deduce the approximate short-term time-base stability needed in a DVI. In the "mark 1" instrument, 0.5 ns inaccuracy (1/100*th* of a resolution cell) over a 4 μs trigger period corresponds to a time-base stability of 1 part in 10^4, a specification exceeded by any crystal oscillator. As well as providing the sampling time-base, the output of the DVI master oscillator is divided in frequency to give the 4 μs internal trigger signal. Long-term master oscillator drift therefore causes timing drift in the trigger signal. An un-stabilized crystal oscillator such as the one in the DVI has a typical frequency stability of 1 part in 10^9 averaged over one day. This gives a timing uncertainty of 86 μs per day, an acceptable figure even in most pulsar observations. When better long-term stability is needed, the trigger signal derived from the DVI master oscillator can be replaced with one derived from an atomic frequency standard.

Finally, dynamic errors in the analog section of the A/D converter need to be considered. Large, high frequency input signal excursions may drive input buffers and on-chip comparators into slew rate or saturation limits, producing an overall converter performance which is much inferior to that predicted on the basis of low frequency, small signal specifications.

The DVI use of an A/D converter is unusual in that Nyquist sampling is not employed. Instead, a 20 MHz bandwidth is digitized at 25 mega-samples/second (Msps) in the prototype system. A TRW TDC1021J four-bit flash A/D converter was chosen since published specifications

suggest that the device should be satisfactory in a DVI role. The maximum conversion rate is 30 Msps and linearity is specified as ± 1/4 LSB. Aperture jitter is 30 ps and communication with the manufacturer suggests that the aperture time (the time taken to latch the latching analog comparators) is about 100 ps. The sampling time offset is typically 10 ns and, because of the method used to transfer data to the output latch, there is a digital output delay of at least one clock pulse period. The A/D converter can be link programmed to produce binary or two's complement code in normal or inverted form. The analog input range is -1 to 0 V but few high frequency analog specifications are available.

Figure 8-8 is a simplified schematic diagram of the A/D conversion module. An LH0033 high speed, unity gain, buffer amplifier drives the A/D converter from a low impedance source, minimising errors due to input bias currents and input capacitance. The diode and associated current limiting resistor prevent the converter input being driven positive by more than a fraction of a volt, in accordance with the manufacturer's recommendation. Local power supply regulation is provided to minimise induced power supply noise and, for the same reason, great care has been taken with the physical layout and power rail by-passing.

Noise is by far the greatest problem in video A/D conversion systems and it is essential that analog and digital grounds be separated except for a single connection at the power supply. In a four-bit converter with an input range of one volt the quantization steps are 66.7 mV apart. The peak-to-peak noise present on TTL signal lines may be up to 0.5 V even in well-designed high speed processors. Allowing digital ground return currents to mix with analog currents is clearly disastrous and every effort must be made to electrically and physically isolate the A/D converter from the remainder of the system.

Figure 8-9 shows the "mark 1" DVI A/D converter module. Separate digital and analog ground-planes have been used and the entire PCB is contained in a metal box, connected at a single point to analog ground. The converter output is connected to the ALU module by a 300 mm ribbon cable which can be un-plugged from the converter end.

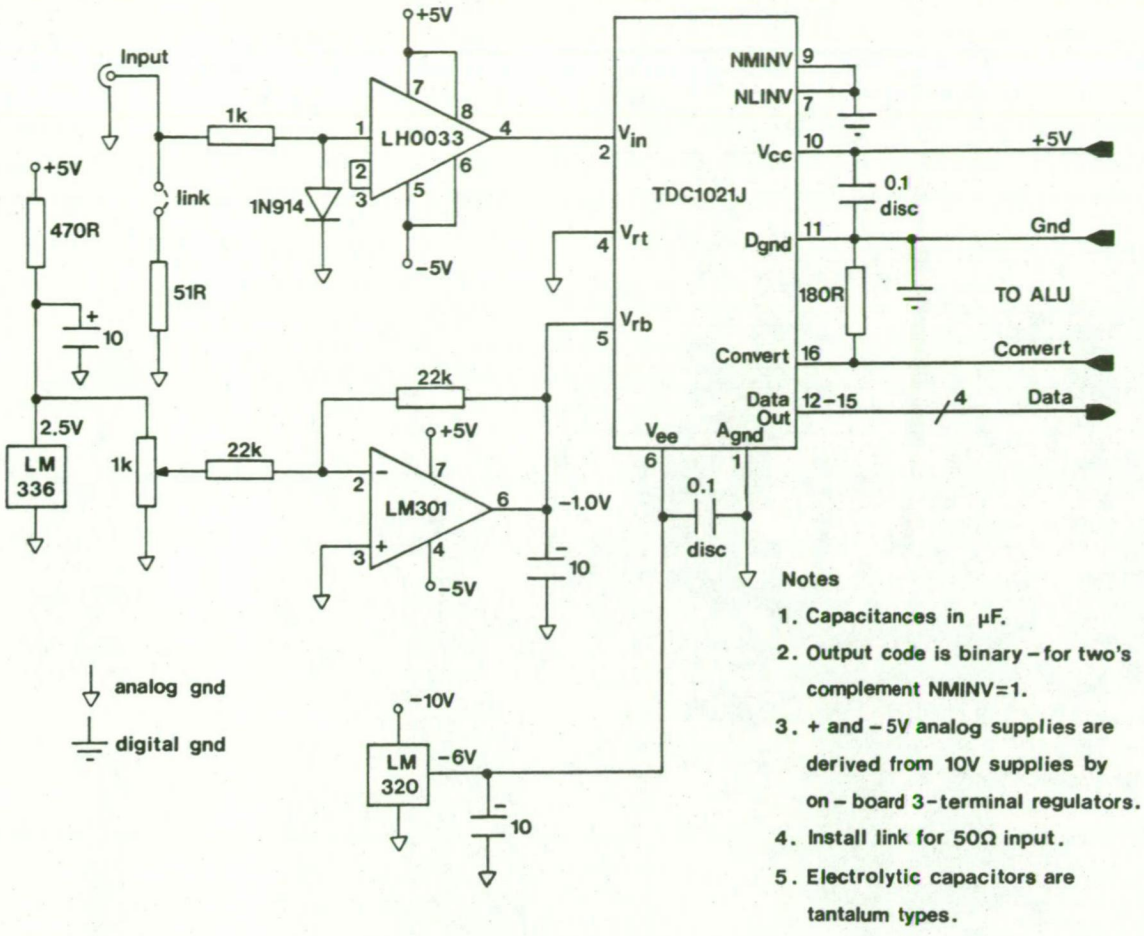


Fig.8-8: Circuit diagram of the four-bit A/D converter module.

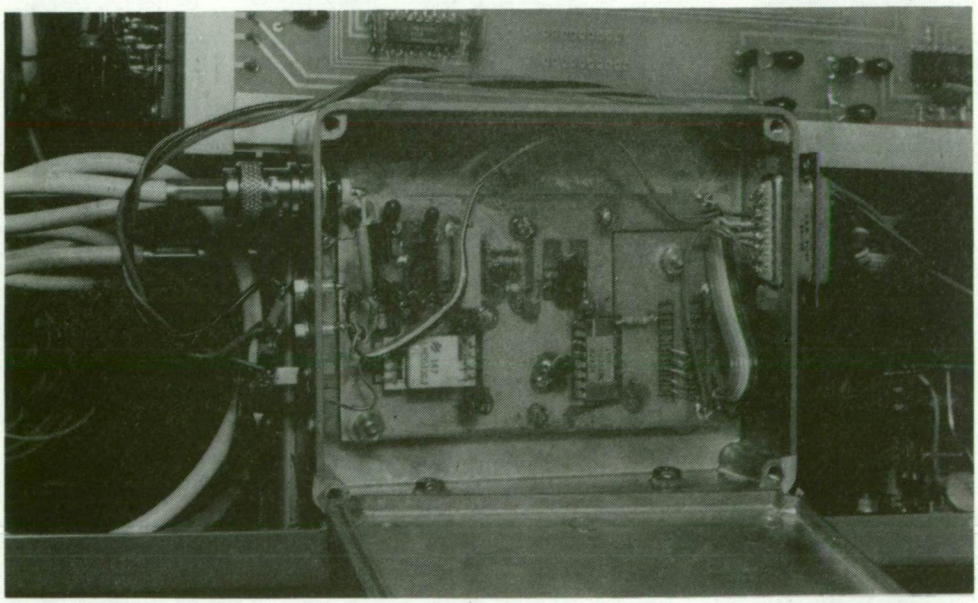


Fig.8-9: Photograph of the A/D converter module. The analog input is on the left; the digital output is via the multi-pin connector on the right. The A/D converter is the IC at the lower right on the PCB.

The cable has alternate earth-active conductor allocation to minimise cross-talk and to provide a low impedance digital ground return between the two modules.

To further reduce the risk of induced noise, the A/D converter output coding has been programmed as straight binary. Thus, for an input range extending ± 0.5 V around -0.5 V, output code 0000 respresents -1.000 V (full-scale negative), code 1000 is -0.5 V (zero) and code 1111 is 0.000 V (full-scale positive). The alternative two's complement code has the advantage of allowing a noisy bipolar input signal to be integrated to a mean of zero but requires a large number of bits to change simultaneously when transitions from positive to negative numbers occur. The large change causes heavy digital power supply currents to flow and can produce a high level of radiated and residual ground loop noise which may adversely affect the video A/D converter. Baseline stripping is invariably performed during data processing and the DVI local integrated spectrum display is unambiguous regardless of data coding. Two's complement code therefore offers no significant practical advantage and since the risk of excess noise is real, the choice of binary code is prudent.

8.6 ARITHMETIC AND LOGIC UNIT

The ALU section of the "mark 1" DVI is shown in block diagram form in figure 8-10. The central elements of the module are a 16-bit adder and a 64 stage, 16-bit circulating accumulator. The adder unit consists of four 74F181 ALU chips and one 74F182 look-ahead carry generator. The worst-case addition time for two 16-bit numbers is less than 15 ns. The 74F181 has both carry generate and propagate signals available so no additional circuitry is necessary when implementing full look-ahead operation.

The "A" inputs of the adder receive data from an output latch (4 × 74F175) via a two way multiplexer (4 × 74F157) and the "B" inputs are presented with data from the A/D converter. The high-order "B" bits are commoned and driven from a 74F08 AND gate. One of the AND gate inputs is permanently high and the other may be either grounded or connected to the A/D converter MSB, depending on whether the ALU is to perform binary or two's complement arithmetic. For reasons discussed in the previous section, binary code is normally used.

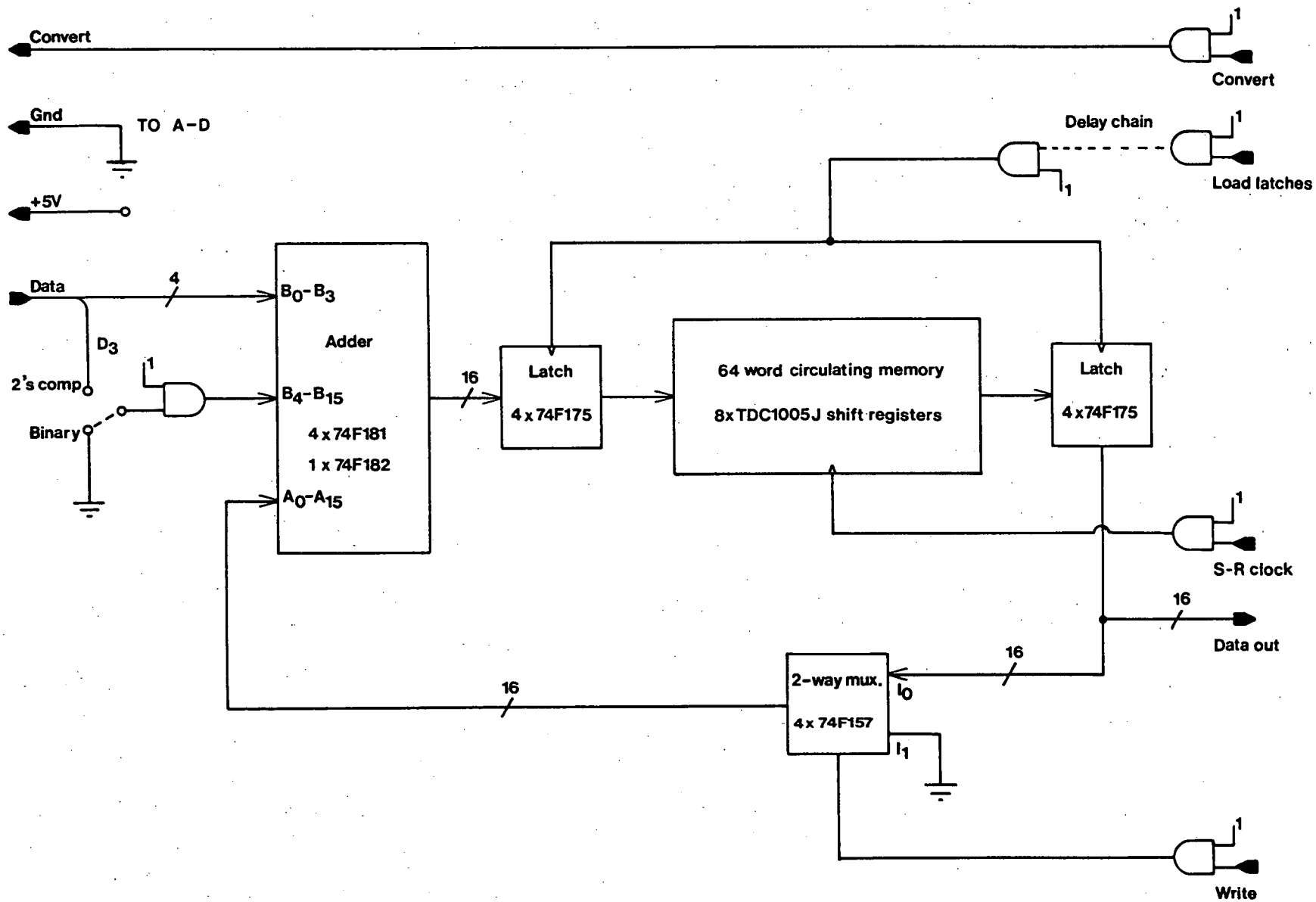


Fig.8-10: Block diagram of the ALU module.

The 64 stage 16-bit circulating memory is actually composed of eight 64 stage, dual shift register IC's. These devices (TRW TDC1005J) are able to be clocked at up to 40 MHz and, like the TDC1021J A/D converters, use advanced bipolar VLSI technology. The IC cost is low enough to make the shift register approach attractive as a method of realizing the memory. If a larger memory length is needed, single 256 stage devices are also available.

A circulating memory based on high speed RAM is feasible but has the disadvantage of requiring both a read and write cycle to take place during the processing interval (typically 40 ns). Even at 15 ns access time, a single RAM is too slow. Blocks of RAM can be commutated but the cost and complexity is increased considerably. A shift register memory has the advantage of requiring only a single clock pulse to read data from the end of memory and write new data to the beginning. The only complication is that, on a nano-second time scale, input data setup and hold times for the VLSI shift registers are quite different from the output data delay time. As well as easing general timing requirements, the two latches shown in figure 8-10 overcome the problem of unequal shift register setup, hold and delay times.

Figure 8-11 is a simplified timing diagram for the ALU. A DVI WRITE state is assumed, so timing for the buffer RAM is also shown. The timing scheme satisfies shift register data setup and hold requirements while still allowing retention of valid spectral data for 64 frequency cells. A master clock frequency of 25 MHz is assumed and representative absolute delay values are shown. In the idle state, data for spectral cell 0 reside at the end of the shift register, so cell 63 data are at the register beginning. An active cycle begins with the commencement of the 66-pulse shift register clock train and the progress of cell data through various ALU stages is shown in the diagram. The shift register first stage output is not available externally but is included for illustrative purposes.

Note that the A/D converter and data latches are clocked continuously and that there is a one clock cycle plus 30 ns A/D converter digital data delay. Data are present at the ALU output during all integrator cycles but are written to the buffer RAM only during a WRITE

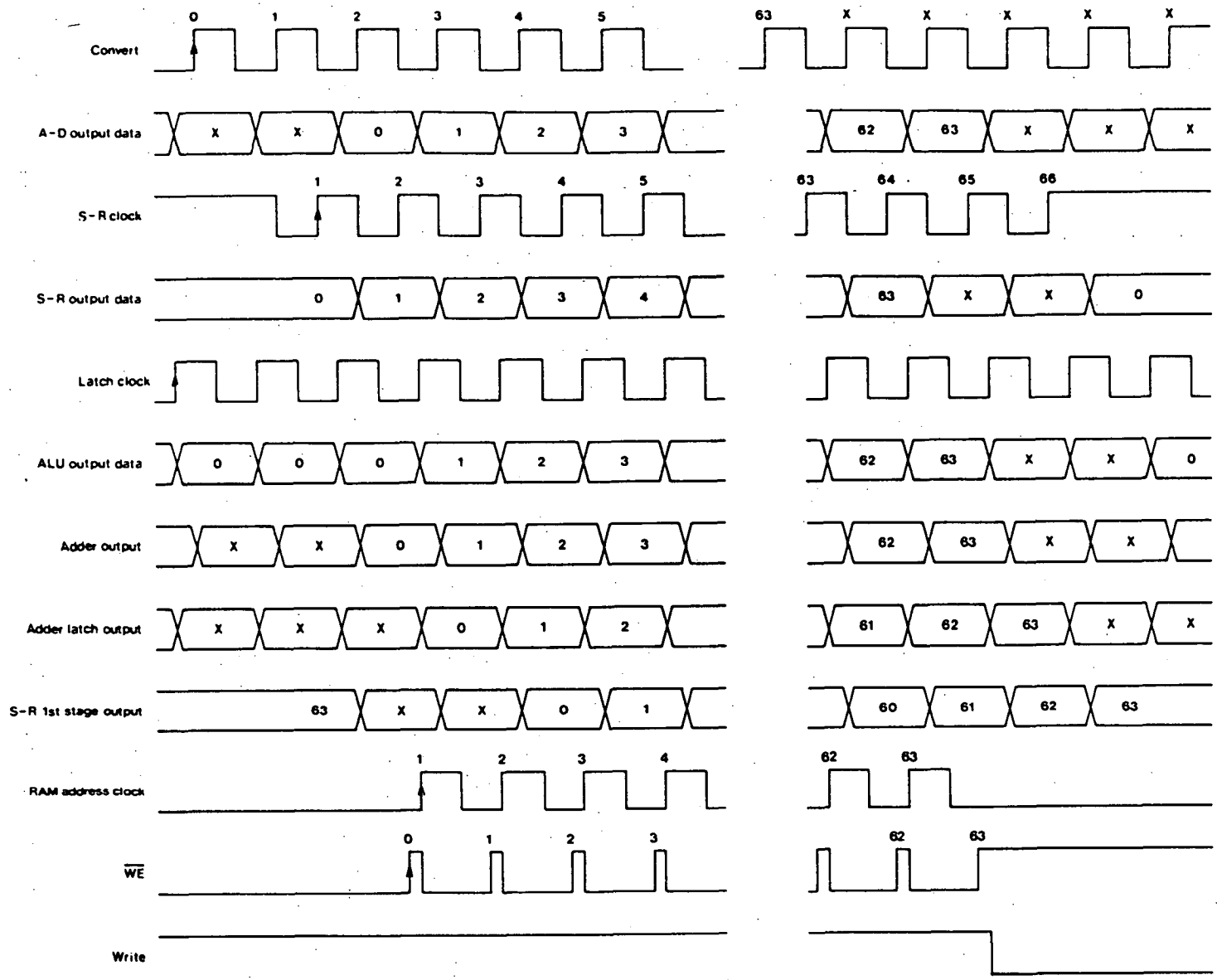


Fig.8-11: Simplified timing diagram for the ALU and buffer RAM modules.

state active cycle. For the duration of a WRITE state, the 74F157 multiplexers open the accumulator-adder loop and force the "A" inputs of the adder to zero, allowing new A/D converter outputs to be written directly into the shift register accumulator to form the old sum for the next (READ state) integrator cycle (section 8.3).

Figure 8-12 shows the ALU circuit board and the RAM board described in the next section. The large power and ground conductors are needed to retain good pulse fidelity and high noise margin. All signals entering the ALU board are buffered and the "load latches" signal can be adjusted in 4 ns increments by selecting one of four delays in a 74F08 AND gate delay chain. The five DVI's so far constructed function with any of three delay settings. However, a selectable delay allows fine tuning of the DVI timing and could be useful if a slow ALU or shift register IC caused excessive data skew.

8.7 OUTPUT BUFFER RAM

The RAM board contains sixteen AMD AM27S07 16 word, four-bit, tri-state output, RAM chips arranged to form a 64 word, 16-bit memory. The RAM chips use ECL technology internally but have on-chip level translators to give TTL compatibility. To give the buffer RAM PCB wider application, provision has been made to extend the board to 128 words with the addition of a further 16 RAM IC's. Figure 8-12 shows an assembled PCB and figure 8-13 is a block diagram of the output buffer. The PCB has been designed with large, extensively interleaved and de-coupled power supply rails. The board accommodates a MSB inverter to convert inverted two's complement numbers to straight binary representations. Links are installed to by-pass the inverter when binary code is produced by the A/D converter and ALU modules.

In the "mark 1" DVI the main requirement of the RAM is that the total time taken to complete a write operation be less than 33 ns (40 ns with 25 MHz sampling). The maximum write enable ($\overline{WE}$) pulse width needed with AM27S07 IC's is 25 ns and no address setup time (prior to $\overline{WE}$) or data hold time (after $\overline{WE}$) is needed for typical devices. The chip enable ($\overline{CE}$) signal is permitted to overlap both the write pulse and data signals, so a write operation can be completed in under 25 ns with typical devices. The memory is read by activating $\overline{CE}$ with $\overline{WE}$ inactive.

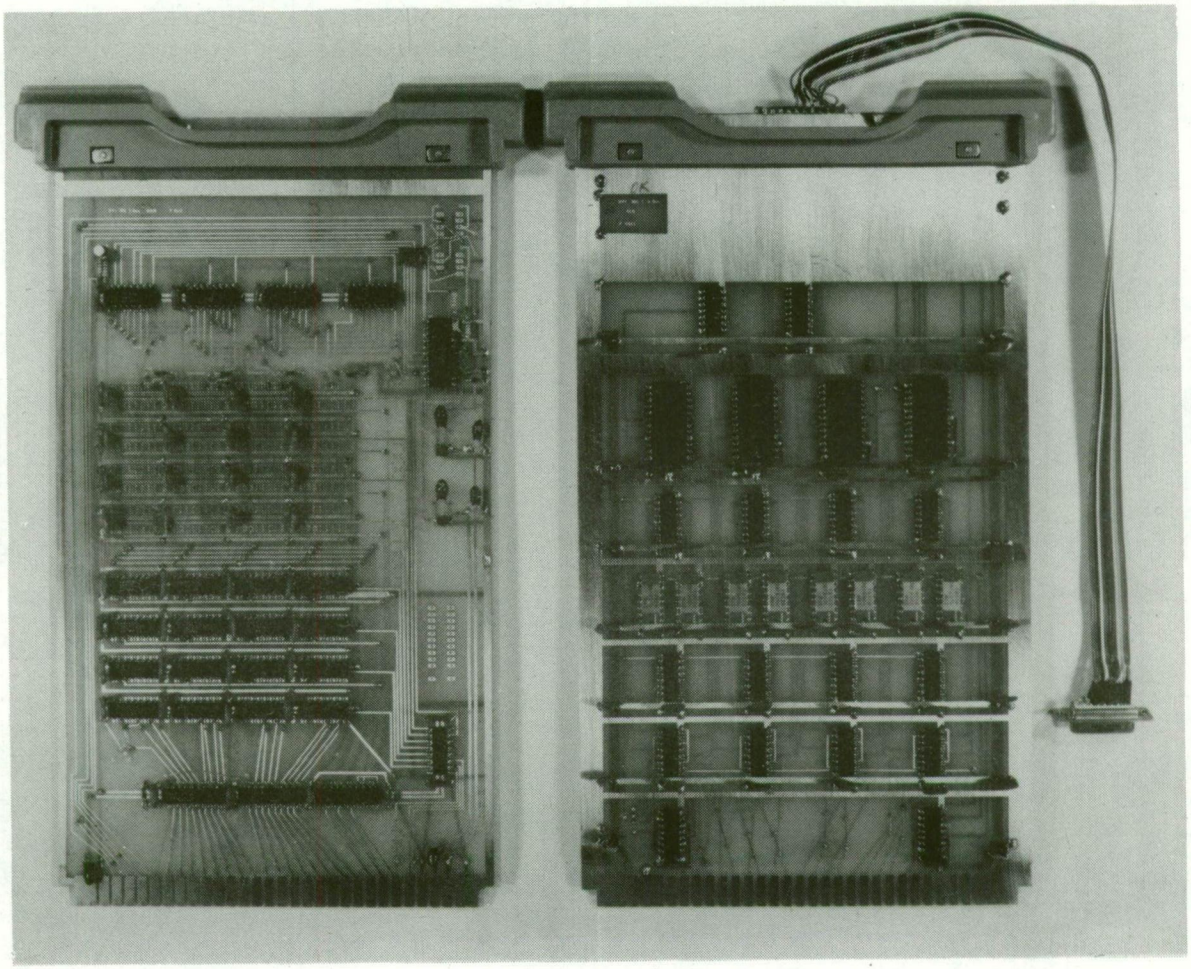


Fig.8-12: Photograph of the buffer RAM (left) and ALU (right) printed circuit boards.

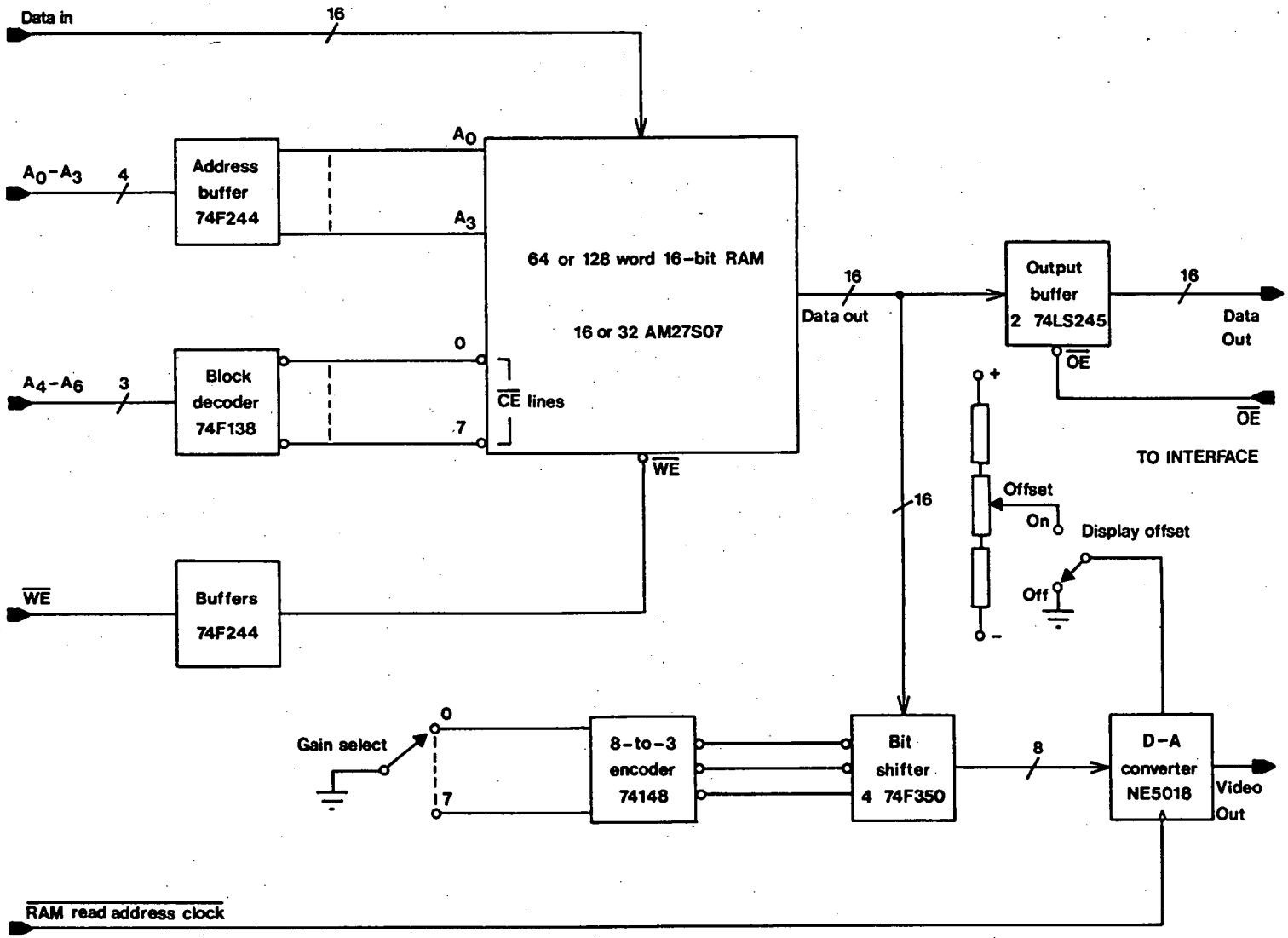


Fig.8-13: Block diagram of the output buffer RAM module.

The 128 word (maximum) memory is divided into eight blocks of 16 words with each block consisting of four AM27S07 IC's. Low-order address bits from the control module RAM address counter select the memory address within a 16 word block and the high-order address bits are decoded in a 74F138 one-of-eight decoder to give a "block select" signal. With a 64 word memory only the first four blocks are ever enabled.

A 74F244 bus driver is used to buffer the $\overline{WE}$ pulse train and the low-order address bits. Data output buffers are 74LS245 tri-state bus transceivers having the "output enable" ($\overline{OE}$) line available at the PCB edge connector. The buffer RAM boards of all four DVI's are connected to a common output data bus and by activating the appropriate "output enable", data from a given DVI are placed on the bus. The individual outputs are usually selected by a computer interface.

As well as the memory and digital output circuitry, the buffer RAM PCB contains a bit shifter and an 8-bit D/A converter. These devices form part of a selectable-gain visual display unit and provide the Y-axis signal for a CRO display of integrated spectral power versus frequency. The bit shifter consists of four 74F350 IC's. In response to a 3-bit control code sent from an eight-to-three encoder (74148) mounted behind the gain selector rotary switch, the shifter moves data by up to seven binary places before presenting it to the D/A converter. The digital output of the buffer memory is not affected in any way by the gain control and all sixteen data bits are transmitted to the data acquisition computer.

The D/A converter is a Signetics NE5018 and has the stages necessary to give a programmable-range output in a single IC. The inverting input of an internal buffer amplifier is accessible and, if the front panel display offset switch is ON, a d.c. offset derived from a ten-turn potentiometer is summed with the D/A converter output. During a DVI READ state, data are clocked into an internal D/A latch on the falling edge of the RAM address counter clock. The address changes on the rising edge, so RAM data are stable before the video display output changes.

8.8 CONTROL SECTION

The control board supplies all the digital signals necessary to operate the other modules and is functionally the most complex DVI sub-section. A simplified block diagram of the control section is shown in figure 8-14 and the timing relationship of major DVI waveforms has been illustrated in figure 8-11. Figure 8-14 is intended as a functional diagram and does not include details of special design precautions taken to prevent race conditions and unwanted states in the high speed logic system. Similarly, details of start-up sequencing are omitted.

In current versions of the "mark 1" instrument, all DVI timing is referenced to a 25 MHz crystal oscillator. The crystal is a high quality type but is not thermally stabilized (section 8.5). The oscillator and associated buffer stages are a 74F04 inverter package.

The trigger pulse generator derives the SAW spectrometer trigger pulse from the DVI master clock square-wave. The generator input may come from either a 74LS390 divide-by-100 IC (internal trigger) or from the external trigger input on the DVI front panel. The output trigger pulse width is half the master clock period (20 ns with 25 MHz sampling) and is synchronized precisely with the clock (figure 8-3). The main components in the trigger generator are 74F74 "D" flip-flops.

The operation of the DVI in "pulsar" mode is detailed fully in section 10.2. However, figure 8-14 indicates that when internal trigger mode is selected, DVI triggering can be prevented by the presence of an externally-supplied reset signal at the 74LS390 divider. This facility allows the DVI operation to be synchronized by a pulsar timing unit.

After the spectrometer is triggered the DVI absolute delay generator (or counter) begins counting master clock pulses. When a programmed number of pulses is reached, the delay counter enables the main DVI control signal generation circuitry and re-arms itself in readiness for the next trigger pulse. Two 74F191 pre-loadable BCD down-counters form the delay generator and give a maximum delay of 10.2 μs with a 25 MHz clock. Links are used to program the delay but DIL switches can be substituted if frequent adjustment is necessary.

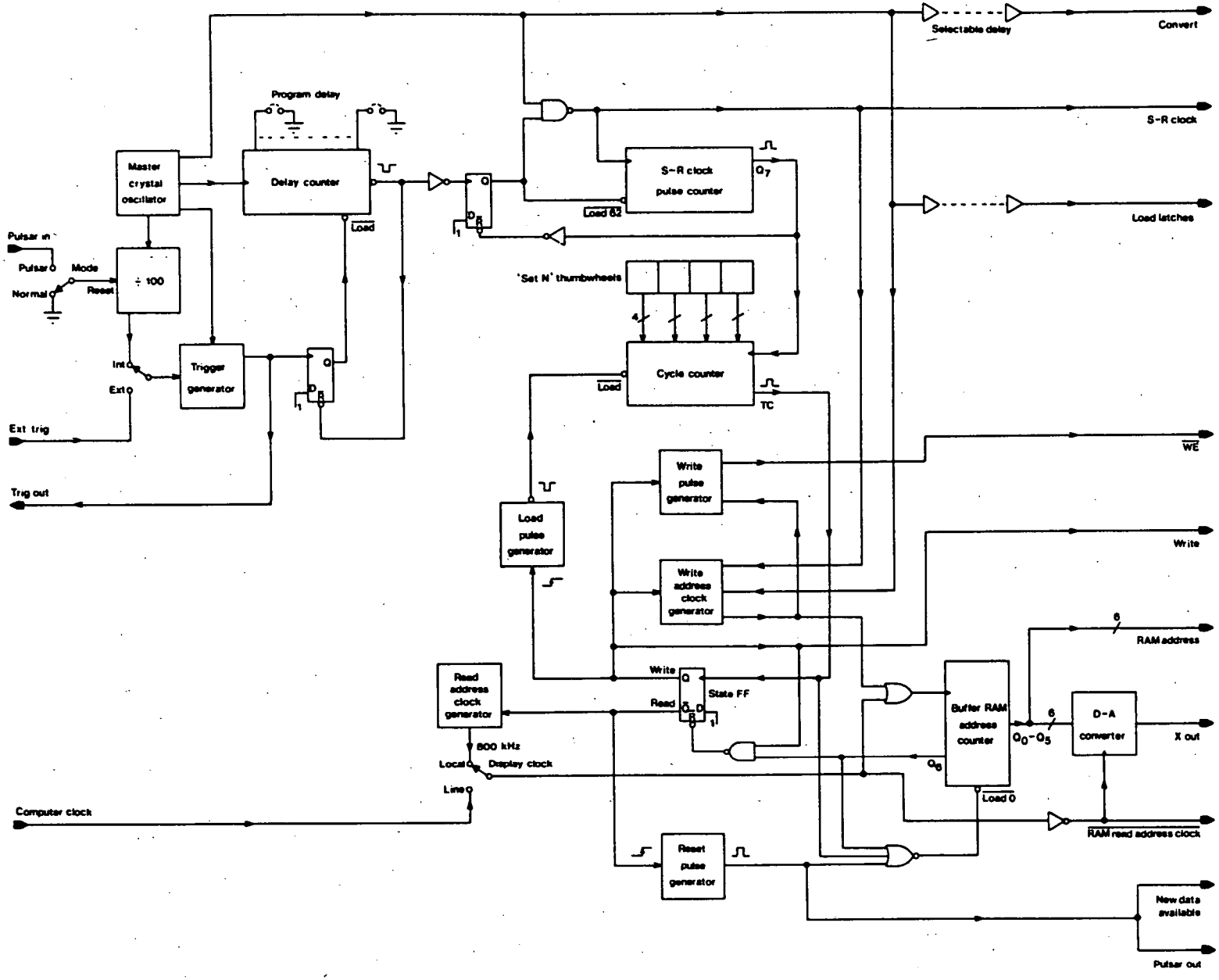


Fig.8-14: Simplified diagram of the control module.

In the prototype system the spectrum analyser trigger interval is 4 μs and the SAW processor delay is about 6.6 μs measured to the beginning of the full-compression window (figure 7-2). The appropriate time interval to program into the delay generator is therefore 6.6 - 4 or 2.6 μs.

The shift register clock generator produces a train of 66 clock pulses every DVI cycle. Each WRITE state cycle, the buffer RAM write address clock generator produces a pulse train offset in time from the shift register clock (figure 8-11). Buffer RAM $\overline{WE}$ pulses are also produced during a WRITE state cycle. The shift register clock is synthesized by combining the master clock with gating signals produced by the delay generator and associated circuitry. The RAM write address clock is derived by combining the shift register clock with hold-off signals and with the master clock. The $\overline{WE}$ pulse train is synthesized by processing the RAM write address clock in a write pulse generator which consists of two 74F74 flip-flops operating in a continuous set-reset loop.

Four 74LS190 pre-loadable BCD counters in a down-count configuration are used as the basis of the DVI cycle counter. During the normal integration process the counter counts the number of complete shift register clock pulse trains that have been produced and, when terminal count (zero) is reached, sets the DVI state flip-flop to WRITE. At the beginning of a WRITE state the counter is loaded with the number contained in the "Set N" thumbwheel switches. At the end of a WRITE state (determined by the RAM address counter reaching its maximum count) the state flip-flop is reset to READ, a "new data available" pulse produced, the RAM address counter reset in readiness for a read operation, and the cycle counting process continued until terminal count occurs again. The cycle counter counts the WRITE state cycle because data from the A/D converter are read into the ALU circulating memory during this cycle.

The buffer RAM address counter consists of two 74F191 counters operating in an up-count mode. The address is taken from the counter "Q" outputs. During a WRITE state cycle the counter is clocked by the RAM write address clock. For local readout of the buffer RAM, the READ state clock comes from a free-running TTL oscillator designed

around a 74LS13 Schmitt trigger. When computer (line) buffer readout is selected, the internal oscillator is disabled and READ state address clock pulses are sent from the computer interface. A Signetics NE5018 D/A converter is connected to the address counter outputs and during the READ state provides a ±5 V ramp for use as an X-axis deflection signal in a CRT integrated spectrum display.

All signals leaving the control board are buffered and in some cases a buffer delay chain is used to allow precise timing adjustment. Input conditioning is also employed. In particular, the main DVI trigger input signal is extensively processed on entry to the trigger generator, ensuring minimal jitter on the SAW spectrometer trigger pulse.

8.9 DVI TESTING AND EVALUATION

The DVI is amenable to most diagnostic techniques applicable to high speed logic systems. However, the high clock frequency means that conventional logic analysers cannot be used to check the instrument operation. If a logic analyser with a sampling capability in excess of 200 MHz is available it can be used as an alignment and trouble-shooting aid although the shift register circulating memory and an easily opened memory-adder loop facilitate testing with only a high speed oscilloscope. Many performance checks are simplified if a triggerable pulse generator is also available.

The overall DVI operation can be checked by integrating input signals synchronized to the master clock. These signals can be produced by triggering a pulse generator from the spectrometer trigger ("trigger out") pulse. The generator pulse delay, pulse width and output offset are adjusted to produce a pulse falling within the DVI sampling window, covering the desired number of cells and positioned at the required level. Moving the generator d.c. offset over the A/D converter input range and adjusting the "Set N" thumbwheel switches should produce the expected changes on the integrator X-Y display.

As well as rectangular pulses, input signals having smooth transitions should be used since fault conditions such as excessive A/D converter dither due to induced noise are more obvious when continuous waveforms are viewed. Transients produced by simple RLC networks excited by a pulse generator often make suitable test waveforms.

Figure 8-15(a) shows a damped sine wave obtained by repetitively impulsing a parallel tuned circuit with a delayed version of the DVI "trigger out" signal. In figure 8-15(b) the DVI output is shown. Note the effect of quantizing and also the clean transitions between quantizing levels.

The results of an extensive program of digital and analog tests indicate correct DVI operation. As well as scaling, linearity and noise checks, an interesting test has been conducted to assess the high frequency (dynamic) performance of the TDC1021J A/D converter. High frequency evaluations are difficult in practice, although the tests described by Peetz et al. (1982) are very effective if a high quality synthesized RF signal generator is available. A suitable synthesizer was not available during DVI commissioning, so one of the tests outlined by Peetz et al. was adapted for use with a stable and spectrally pure analog signal generator. The resulting procedure still yields a large amount of information about the dynamic operation of the A/D converter.

The method relies on the fact that digitizing a sine wave offset slightly in frequency from the converter sampling frequency (or its harmonics or sub-harmonics) produces a beat waveform which appears in the DVI buffer RAM. Provided the beat frequency is not high enough to tax the video D/A converter, the waveform read from the buffer is a reconstructed version of the input signal as encoded by the A/D converter. If the input to the A/D converter is a clean sine wave, distortion effects such as slew rate limiting or dynamic code omission are clearly visible. The input frequency must be a multiple of 250 kHz (the DVI internal trigger frequency with a 25 MHz master clock) so that coherent addition between sampling frames occurs. If the beat pattern is established and the DVI triggering disabled by selecting external trigger mode with no external source present, the last beat waveform is frozen in the buffer and remains stationary for examination.

Figure 8-16 shows the beat pattern obtained when digitizing a full-scale 25.750 MHz sine wave input. The output coding is monotonic but there is evidence of dynamic output code omission, analog saturation effects and slew-rate limiting. Tests around 12.5 MHz indicate similar defects although the distortion is less severe.

(a)

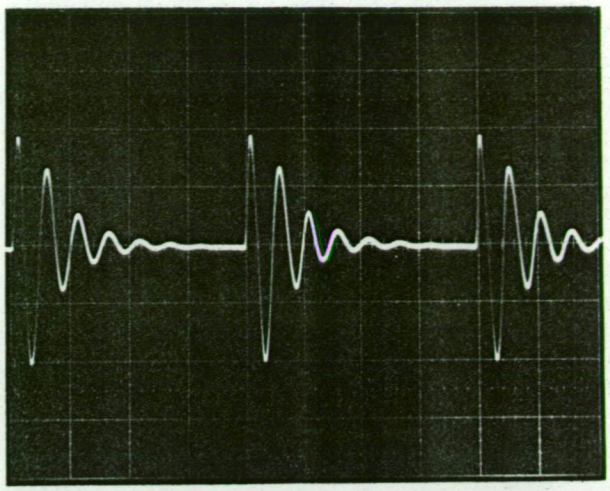


Fig.8-15: Example of a DVI test. In (a) the damped sine wave input signal is shown. The DVI output is shown in (b). The scales in (a) are 1 μs/div. and 200 mV/div.

(b)

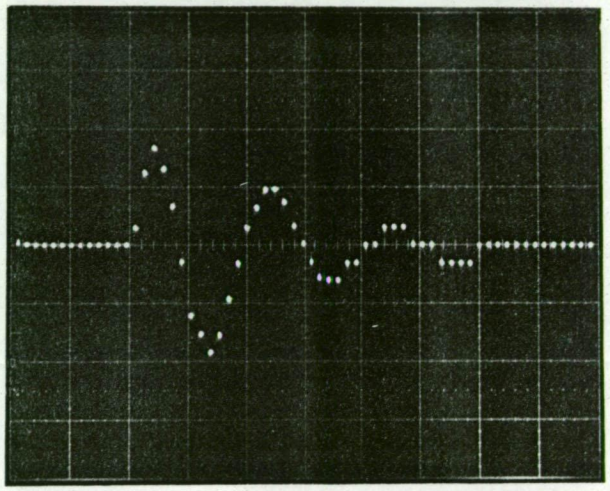
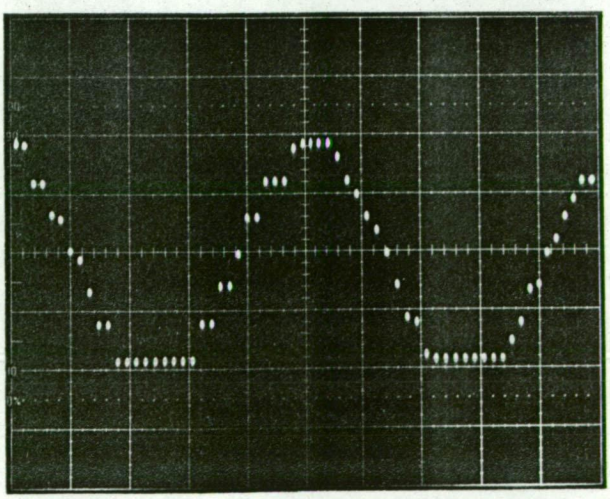


Fig.8-16: DVI output with a full-scale 25.750 MHz sine wave input. The input signal is symmetrical around -0.50V. Disregarding the dynamic code omission, 12 quantizing levels are evident, indicating a fall of about 2 dB in the large-signal frequency response. Increasing the input level further produces gross saturation and slew rate distortion in the beat waveform.



The overall result is an A/D converter which departs quite significantly from an ideal four-bit device. All four A/D converter modules show the same general behaviour. The effect of dynamic code omission when integrating noisy signals is not easy to predict. However, it would be surprising if the A/D converter performance were any better than that of an ideal three-bit converter. Use of the beat test with 50, 75, and 100 MHz input signals suggests that the small-signal frequency response of the A/D converter module is 3 dB down at around 45 MHz, with the response above the break-point being approximately fifth order.

CHAPTER 8 - REFERENCES

Alldritt, M., Jones, R., Oliver, C.J. and Vaughan, J.M.: The processing of digital signals by a surface acoustic wave spectrum analyser. J. Phys. E: Sci. Instrum., Vol.11, 1978, pp.116-119.

Bird, S.C. and Folchi, J.A.: Time base requirements for a waveform recorder. Hewlett-Packard Journal, Nov. 1982, pp.29-34.

Peetz, B.E., Muto, A.S. and Neil, M.N.: Measuring waveform recorder performance. Hewlett-Packard Journal, Nov. 1982, pp.21-29.

Racal-MESL Microwave Ltd.: W1740 series spectrum analyser specifications. Edinburgh, 1982.

Swartzlander, E.: VLSI architecture - in "Very large scale integration" (D.F. Barbe ed.). Ch.6, Springer-Verlag, 1982.

Zuch, E.L.: Video analog to digital conversion - in "Data acquisition and conversion handbook" (E.L. Zuch ed.). Datel-Intersil, 1979, pp.74-79 and 106-107.

CHAPTER 9 SAW COHERENT RADIOMETER PERFORMANCE

9.1 INTRODUCTION

In this chapter the results of performance checks conducted on the prototype SAW coherent radiometer (IF processor and DVI's) are reported. Tests were made with the DVI's interfaced to a PDP11/34 computer system having graphics and plotting facilities. All investigations were conducted using the direct-coupled multiplier module (figure 6-11). This module incorporates a Schottky diode clipper, allowing high-level noise to be applied to the TDC1021J A/D converter without driving the analog stages of the converter into saturation on peak noise excursions.

9.2 RESULTS

The effectiveness of a spectrum integrator as an accessory to a SAW spectrometer is illustrated in figure 9-1. White noise was added to a weak CW signal at the spectrum analyser input. Although almost invisible in the fast noise at the analyser detector output (there is some photographic integration), the signal becomes progressively more obvious as the number of integrations is increased and the r.m.s. background variation decreases as $1/\sqrt{N}$. The spectrometer response is about 3 dB down at the band edges and the passband shape is readily apparent.

The linearity of the SAW radiometer intensity channels is shown in figure 9-2. Characteristics of ideal two and three-bit A/D converters when quantizing exponential noise are also shown. The characteristics are taken from figure 5-3(a), the horizontal axis being offset to bring the arbitrary power scales into coincidence. The IF stages of the broadband receiver described in section 10.4 were used in conjunction with an external noise source to produce test signals, so the input power scale labels in figure 9-2 represent receiver attenuator settings.

Figure 9-3 illustrates the noise-to-signal performance of the radiometer as a function of input level. To produce the graph the DVI was set to N = 1000 integrations so, if the noise-to-signal ratio at the detector output is unity, the expected NSR is $1/\sqrt{1000}$ or -15.00 dB.

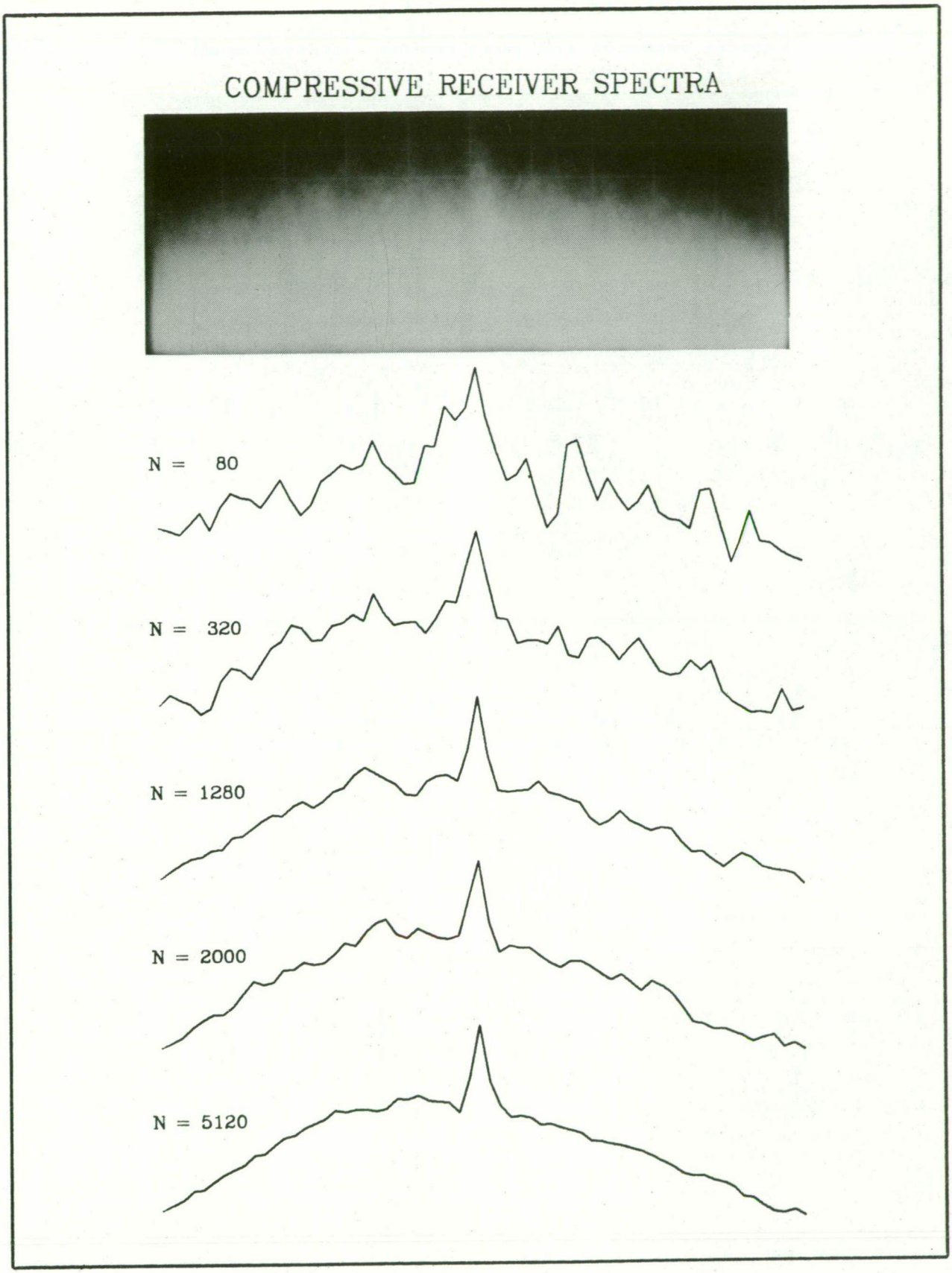


Fig.9-1: Spectra obtained using the "mark 1" DVI in conjunction with the prototype SAW processor. The video output of one intensity channel is shown at the top and spectra obtained using various numbers of post-detection integrations are shown in the lower traces. As N increases the background fluctuation decreases, making the CW signal more visible.

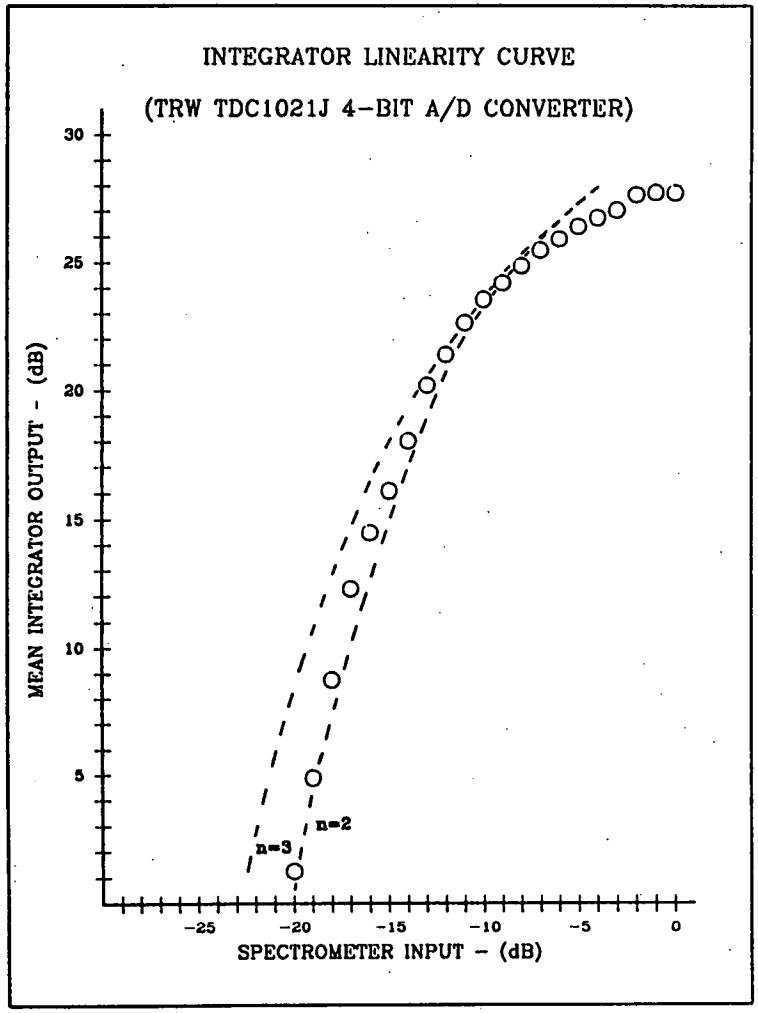


Fig.9-2: Linearity of the SAW coherent radiometer intensity channels. Curves for integrators having ideal two and three-bit quantizers are also shown.

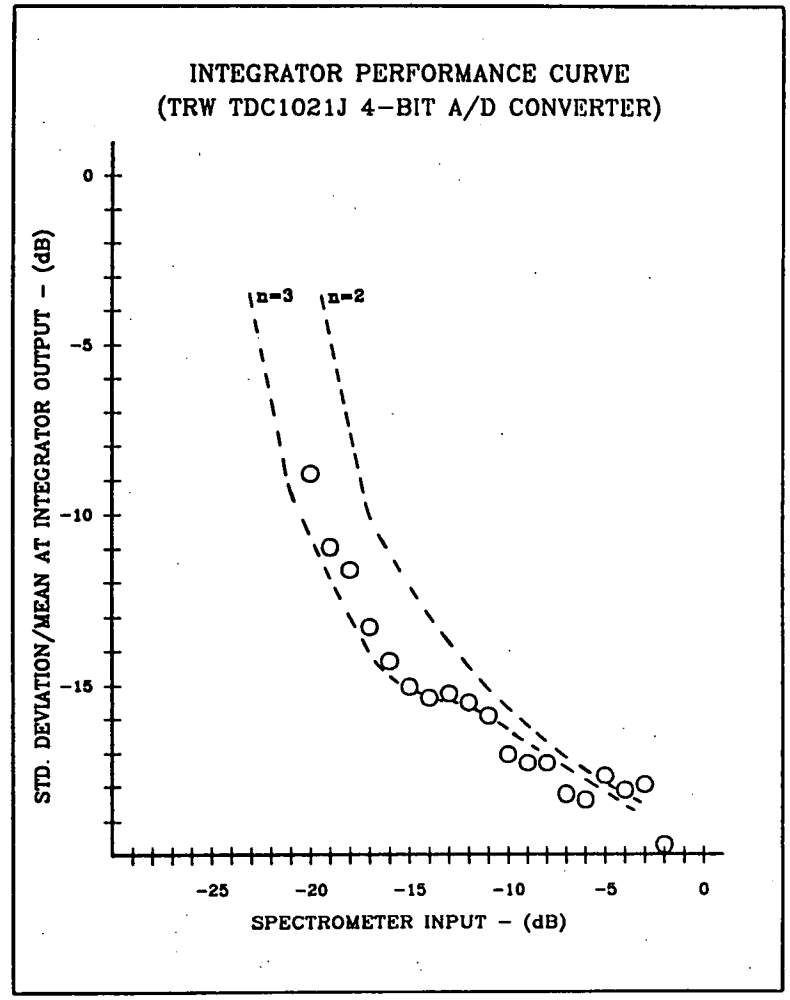


Fig.9-3: Noise-to-signal ratio performance of the SAW coherent radiometer intensity channels. As in figure 9-2, performance curves for integrators having ideal two and three-bit quantizers are given.

Over the optimum input range the NSR is close to -15.5 dB, the estimated measurement error being ± 0.2 dB. The NSR characteristics for ideal two and three-bit quantizers have been transferred from figure 5-4(a). The idealized curves have been shifted horizontally by the same offset used in figure 9-2 and vertically to bring the plateau region of the three-bit curve close to the flat region of the actual curve.

The linearity and NSR checks indicate that the general behaviour of the signal processing system is as expected, especially in view of the poor performance of the four-bit A/D converter (figure 8-16). At optimum input levels the A/D converter approximates a three-bit quantizer while at lower levels the performance deteriorates towards that of a two-bit device. The sharper than predicted saturation characteristic at very high input levels is due mainly to the onset of overload in the SAW spectrometer.

The curves of figure 5-4(a) indicate that the quantizing noise contribution from a three-bit A/D converter degrades the SNR by about 0.5 dB at optimum input levels. Thus, figure 9-3 suggests that an ideal integrator would give a NSR ~ -16 dB after 1000 integrations, corresponding to a NSR ~ -1 dB or 0.8 at the quantizer input. To confirm the NSR behaviour, an optimum input level was set and N (the number of post-detection integrations) varied. Figure 9-4 shows the NSR as a function of $\log_{10}N$. The slope of the line of best fit is -5.01, indicating that a $1/\sqrt{N}$ noise reduction mechanism is in operation. When the line is extrapolated to N = 1 (log N = 0) the NSR is -0.4 dB. If a 0.5 dB quantizing degradation is assumed, the conclusion is as before: the NSR at the quantizer input is about -1 dB. This is discussed in the next section.

Evaluation of the radiometric efficiency of the system requires a knowledge of the bandpass characteristic of the SAW analyser frequency resolution cell. The power response curve was obtained by examining the detected output of a cell in the centre of one compressive receiver passband as a synthesized CW signal generator was stepped across the analyser input band. The resulting response is shown in figure 9-5. Video noise from the direct-coupled

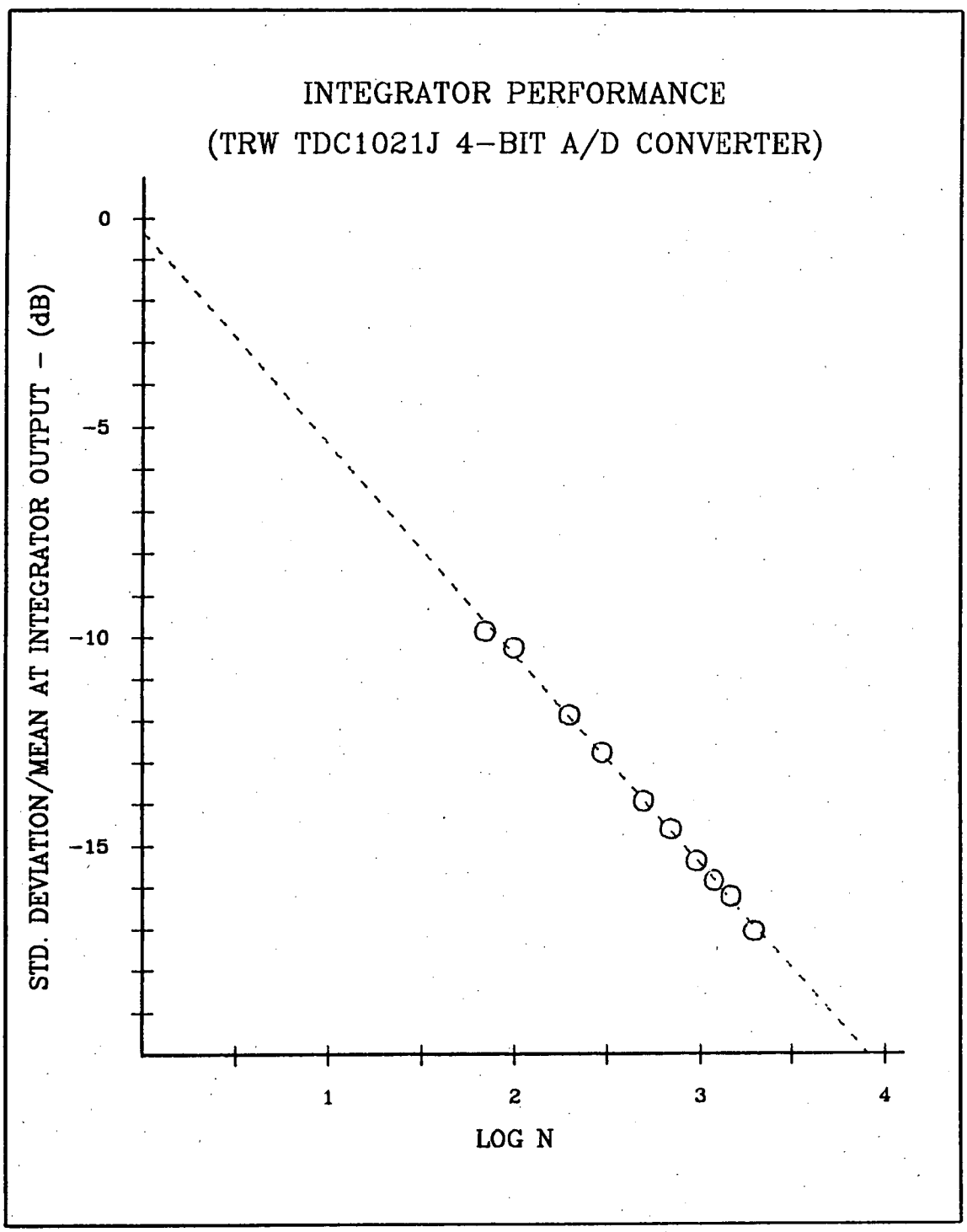


Fig.9-4: Output noise-to-signal ratio of the SAW coherent radiometer intensity channels plotted as a function of $\log_{10}N$ (N is the number of post-detection integrations). A line of best fit is also shown.

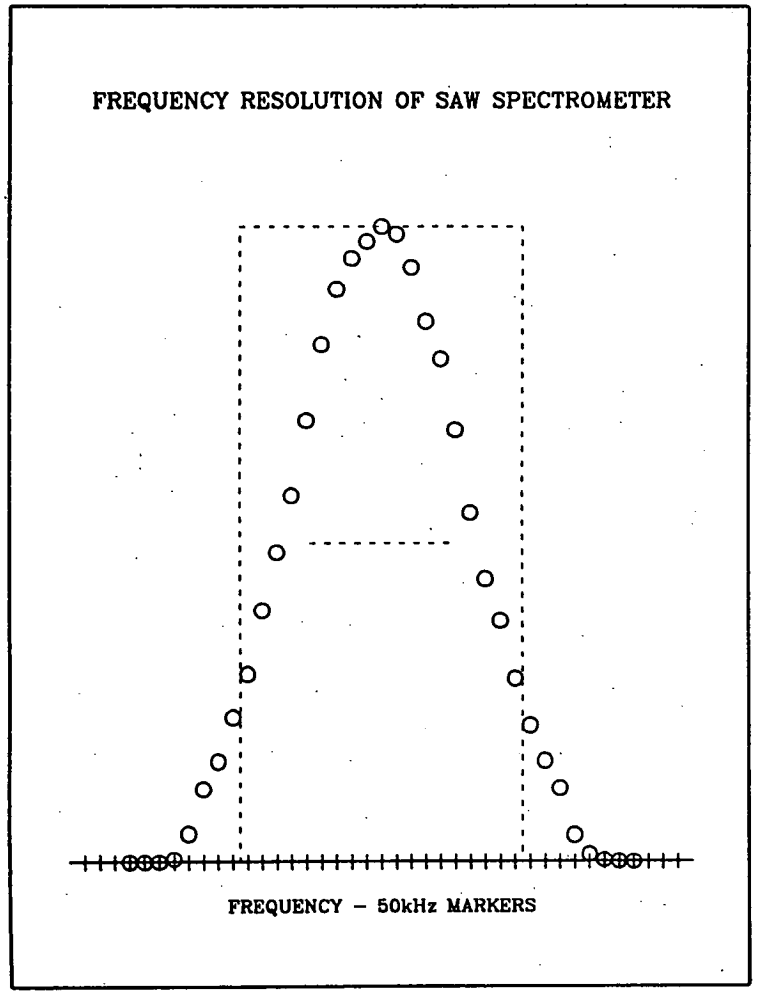


Fig.9-5: Resolution cell shape of the SAW radiometer. The rectangle indicates the equivalent radiometer bandwidth and the horizontal bar shows the half-power (-3 dB) height. The radiometer bandwidth is 950 kHz and the half-power bandwidth is 650 kHz.

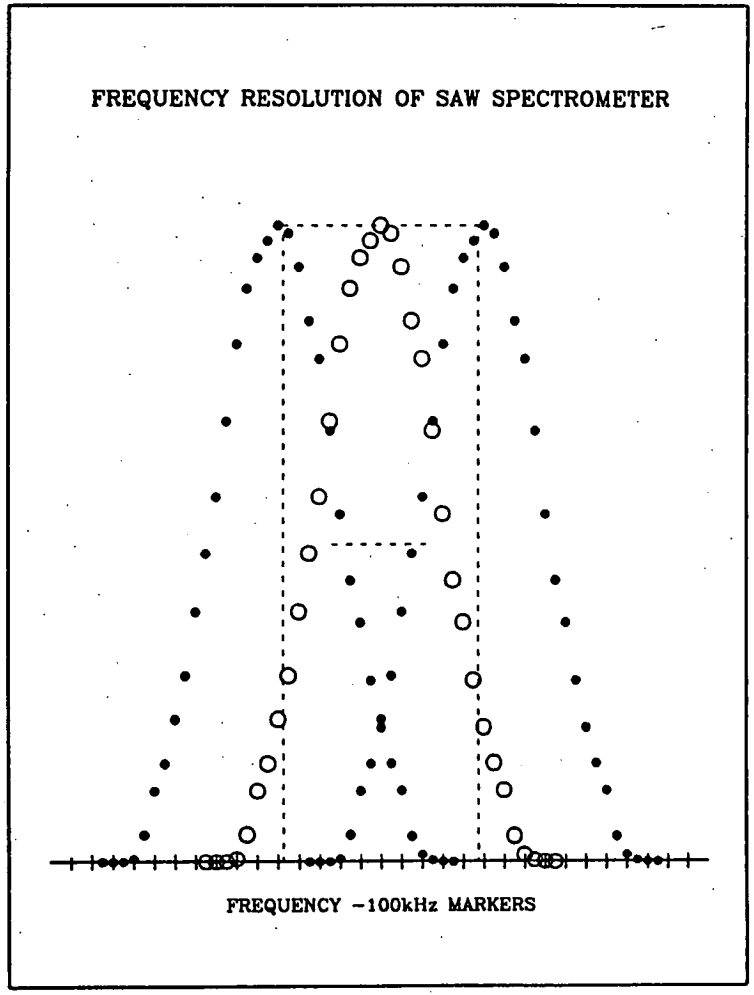


Fig.9-6: SAW radiometer spectral sampling characteristic. Adjacent cells overlap at the -1.3 dB points.

detector module dithers the A/D converter sufficiently to minimize quantizing discontinuities but sidelobes fall well below the threshold needed for effective quantization .

Before evaluating the noise efficiency of the radiometer, it is appropriate to examine the spectral sampling characteristic obtained with the instrument. Figure 9-6 shows the response of adjacent resolution cells. Cell overlap occurs at the -1.3 dB point, giving slightly closer sampling than the usual filterbank arrangement.

The -3 dB width of the cell (denoted by Δf) is 650 kHz and in figures 9-5 and 9-6 the half-power (-3 dB) height is indicated by a horizontal bar. The noise bandwidth is

$$\Delta f_n = \frac{1}{G_{max}} \int_0^\infty G(f)\, df \ , \qquad 9.2.1$$

where G_{max} is the maximum value of the power response, G(f). For figure 9-5 Δf_n = 688 kHz. The radiometer bandwidth is defined as

$$\Delta f_r = \frac{\left[\int_0^\infty G(f)df\right]^2}{\int_0^\infty G^2(f)df} \qquad 9.2.2$$

(Tiuri, 1964). The response of figure 9-5 gives Δf_r = 950 kHz, indicated by the dotted rectangle. To obtain Δf_n and Δf_r, Simpson's rule was used to integrate numerically the power response of figure 9-5. The cell shape is similar to a Gaussian filter response so $\Delta f_r \sim 1.5\,\Delta f$ is a plausible result. The contribution of sidelobes to the radiometer bandwidth has been estimated at less than 5% for a cell at the passband centre. The estimate was made using a manual numerical integration technique on a photographically enlarged oscillogram.

Using N = 1000 as a working point, the time taken by the DVI to complete an integration is 4 ms. An ideal radiometer with Δf_r = 950 kHz and an integration time of 4 ms would produce

$$NSR = \frac{1}{\sqrt{950\times10^3\times4\times10^{-3}}} = 0.016 \text{ or } -17.9 \text{ dB.}$$

The duty cycle of the SAW spectrometer is 60%, leading to a 1.1 dB increase in the NSR. The expected value of the NSR is therefore -16.8 dB.

The best estimate of the observed value is -15.4 dB, corresponding to 72% radiometer efficiency. If account is taken of a probable 0.5 dB quantizing loss, the efficiency with the ideal integrator would be 81%. These implications of these results are discussed in the next section.

To test the quantized correlators, both broadband receiver channels were used as sources of uncorrelated background noise. A pulsed, variable power, correlated noise component was injected into the two channels. Essentially all the uncorrelated noise originated before the receiver attenuators, so the attenuators allowed adjustment of the total noise power. With no correlated noise, the three receiver attenuation values shown in the correlator linearity plot (figure 9-7) result in the same power at the square-law detectors as equivalent values in figures 9-2 and 9-3.

The correlator linearity is plotted as a function of the ratio of correlated to uncorrelated noise power for the three attenuation settings. Over the linear operating range the dotted lines have the slope of ideal characteristics. Horizontal axis labelling is such that 0 dB corresponds to equal correlated and uncorrelated noise power. The most obvious feature of the graph is the variation in correlation gain with attenuation setting. The system is linear at each setting provided the total power does not overload the SAW spectrometer or drive the Schottky diode clipper at the correlator output into hard limiting. At the other extreme, extension of the test to very lightly correlated signals shows that an adequate overall level is needed to maintain good quantizer performance.

Egau (1981) discusses the sensitivity of various types of correlator and an expression given by him for the ratio of the mean to standard deviation at the output of an analog correlator can be modified to yield the expected NSR in the present instrument. If σ_s and σ_n are the r.m.s. levels of the correlated and uncorrelated components at the correlator inputs, after N post-correlation integrations

$$NSR = \left(\frac{\sigma_n}{\sigma_s}\right)^2 \frac{1}{\sqrt{2N}} \qquad 9.2.3$$

if N is large.

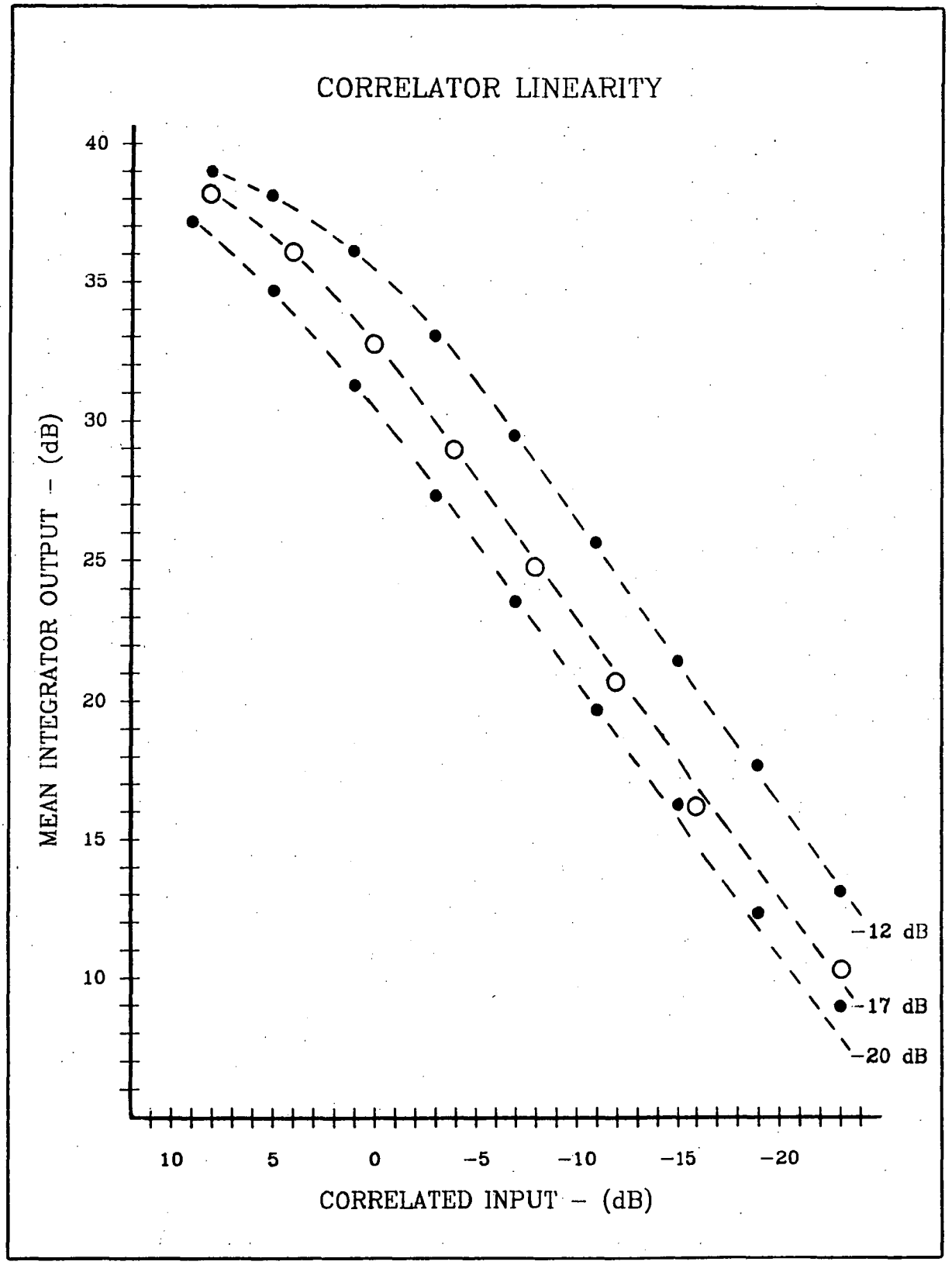


Fig.9-7: Linearity of the SAW coherent radiometer correlation channels plotted as a function of correlated input level for three values of receiver attenuation. Intermediate data points are omitted for clarity.

In the SAW coherent radiometer the quantized correlators exhibit noise-to-signal ratios within 1.5 dB of the expected values over a wide range of receiver attenuator settings. Both linearity and NSR tests indicate that the radiometer correlation channels have greater dynamic range than the square-law detector channels. This is presumably because of quantizer dithering at low levels by uncorrelated background noise.

To demonstrate the effectiveness of the coherent radiometer as a phase measurement system, a delay line was inserted in one receiver output. Vector voltmeter measurements made with the pulsed noise generator replaced temporarily by a tunable CW source led to a predicted phase slope of 9.0^{o}/MHz (4.5^{o}/frequency cell) after compensation for the instrumental phase characteristic (figure 7-12). The characteristic of a signal splitting hybrid gave a small additional offset below 60 MHz, resulting in an extra 2^{o} phase shift at 45 MHz.

With the noise source re-connected, the phase angle

$$\phi = \arctan\left(\frac{V}{U}\right) = \arctan\,(R) \qquad 9.2.4$$

was calculated for each cell and the formal error

$$\sigma_{\phi}^{2} = \frac{R^{2}}{[1+R^{2}]^{2}}\left\{\frac{\sigma_{V}^{2}}{V^{2}} + \frac{\sigma_{U}^{2}}{U^{2}}\right\} \qquad 9.2.5$$

computed explicitly. In these expression V and U are the mean outputs of the quadrature and in-phase correlators after correction for any gain difference between the correlators, and σ_V and σ_U are the on-signal standard deviations.

Figure 9-8 shows the phase characteristic obtained with P_c, the ratio of correlated to uncorrelated noise power, set at -1 dB. The error bars indicate $\pm 2\sigma_{\phi}$ limits and the dotted line is a weighted best-fit line through all points. The magnitude of the line slope is 4.5^{o}/cell, in excellent agreement with the predicted value. An additional low frequency phase shift is evident, with the error at 45 MHz being about 6^{o}, rather more than the 2^{o} predicted. Measurements at P_c = -20 dB give the slope of the best-fit line as 4.92^{o}/cell.

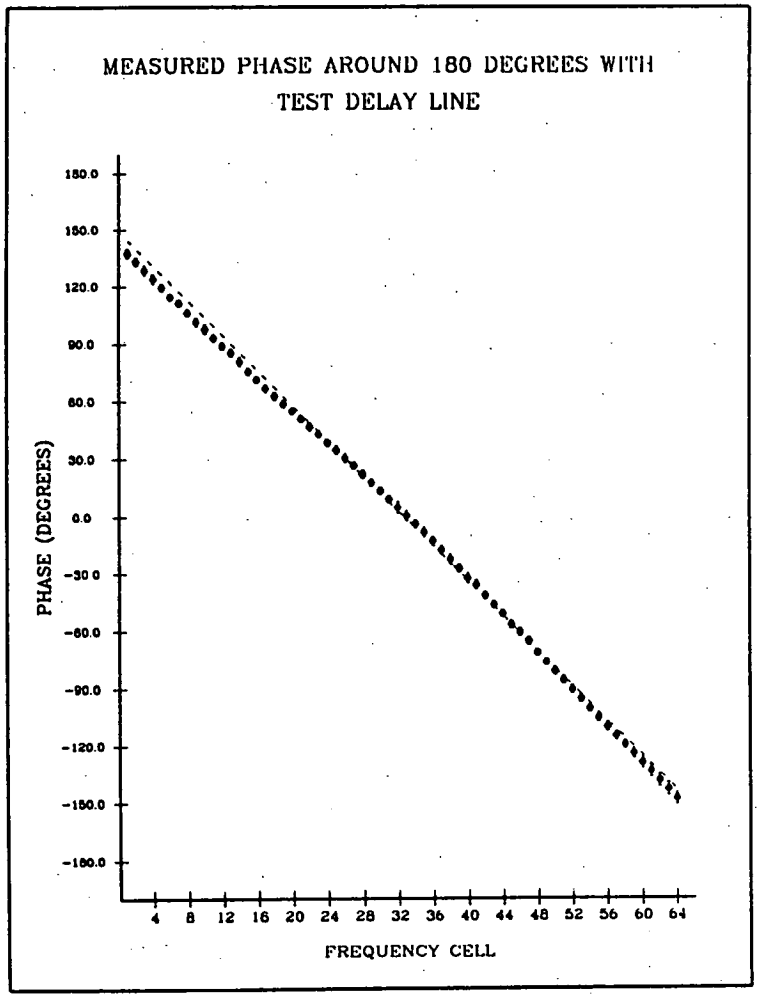


Fig.9-8: Phase-frequency characteristic of a delay line determined using the SAW coherent radiometer. Error bars are ± 2σ and a weighted line of best fit is shown.

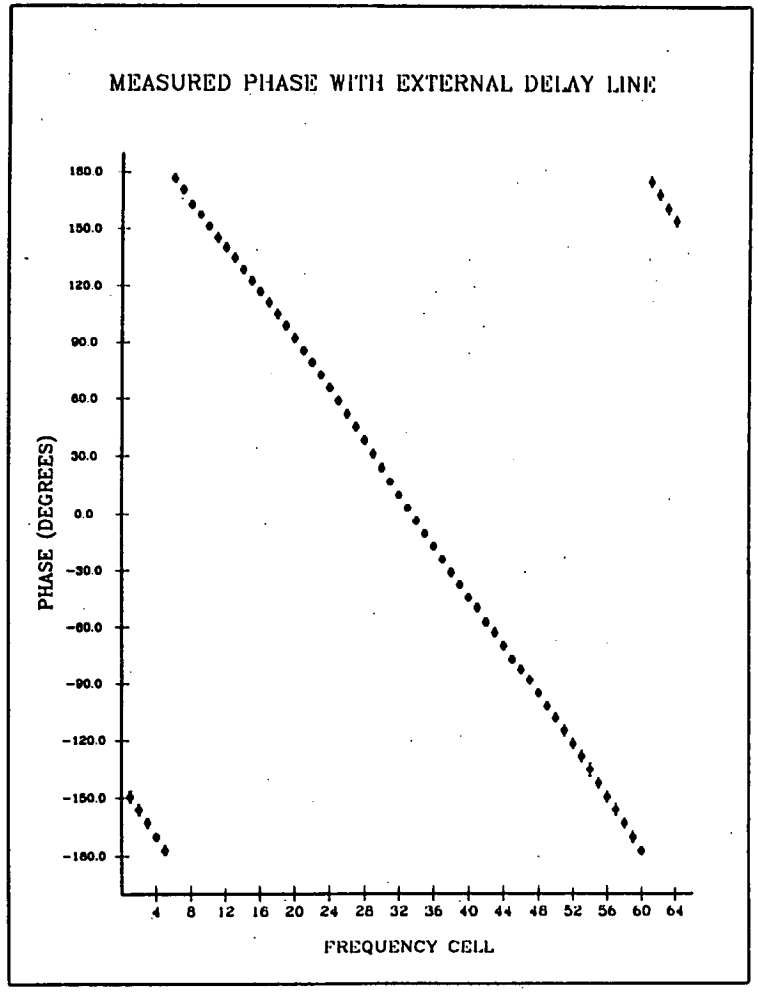


Fig.9-9: Measured phase-frequency characteristic of the extended delay line.

Finally, to verify the continuity of the system around extreme phase values, the length of the external delay line was increased. The resulting phase characteristic (P_C = -1 dB) is shown in figure 9-9. A similar response is obtained at all values of P_C provided signal levels are set correctly.

9.3 CONCLUSION

The results reported in the previous section indicate that the prototype SAW system is a usable scientific instrument. After correction for 60% duty cycle, the measured radiometer efficiency is 72%. Thus, at 100% duty cycle the efficiency should compare favourably with that of one-bit correlators (64%) and typical acousto-optical spectrographs (67% in the unit described by Cole and Milne, 1977). At present, 60% duty cycle operation gives an uncorrected radiometer efficiency of 56%; that is, the sensitivity of the existing system is 56% of that which would be obtained from an ideal continuously operating radiometer having an equivalent bandwidth and integration time. Replacement of the A/D converters with higher resolution devices would improve this figure slightly.

Since the preparation of the majority of this thesis it has become possible to quantify more precisely the expected efficiency of the SAW system. This is done in appendix seven, a note added in proof. In essence, even a continuously operating SAW radiometer exhibits a SNR loss relative to an ideal radiometer. Comparing the sensitivity of the actual SAW instrument with that of an ideal radiometer by simply scaling the ideal SNR by a factor of $\sqrt{d}$ (where d is the fractional duty cycle) therefore creates an artificial discrepancy between expected and measured sensitivity values. When a more exact comparison is made (appendix seven), the agreement between actual and expected values is good. At 56% uncorrected efficiency the prototype radiometer is yielding essentially optimum performance for a 60% duty cycle. Increasing the duty cycle to 100% should therefore give a sensitivity which is close to the maximum theoretical figure (about 70% of that obtainable from a continuously operating ideal radiometer).

In the previous section it was shown that the NSR at the quantizer input is about -1 dB (~0.8), implying that the p.d.f. at this point is not exactly exponential in form. This may be due to the slight spectral smearing effect of the post-detection video filter. The departure is small in radio engineering terms and the simulated integrator curves presented in figures 9-2 and 9-3 are considered useful guides to the expected performance. It is unlikely that the measured deterioration in low-level quantization characteristics is related to the non-exponential p.d.f. since the small amount of extra smoothing should, if anything, improve the system performance slightly.

The results of the correlator tests show that effective quantization is possible in spite of uncertainty about the form of the correlator p.d.f. (section 5.5). The variation of correlation gain with background level makes it necessary to calibrate the system on each source during an observing session. However, this is standard practice and presents no special problems.

Frater (1984) has suggested two mechanisms by which the correlator gain could vary with signal level. Firstly, the self-detection balancing depends to some extent on the input level. Although this effect appears weak in the MM109 multiplier, the peak to r.m.s ratio is quite different for CW and noise inputs. Thus, unbalance on noise peaks could be causing an apparent gain change.

Secondly, it is possible that on-chip gain stages in the MM109 could be affected by the presence of high-level post-correlation RF components. Details of the MM109 internal circuitry are not available but it seems probable that any overload mechanism would produce a decrease in correlation gain at high levels. The opposite is observed, suggesting that the first mechanism is the more likely.

CHAPTER 9 - REFERENCES

Cole, T.W. and Milne, D.K.: An acousto-optical spectrograph for spectral integration. Proc. Astron. Soc. Aust., Vol.3, No.2, Sept. 1977, pp.108-111.

Egau, P.C.: Ph.D. thesis. University of Sydney, 1981.

Frater, R.H.: Private communication, 1984.

Tiuri, M.: Radio astronomy receivers. IEEE Trans. Ant. Prop., Vol.AP-12, No.7, Dec. 1964, pp.930-938.

CHAPTER 10 RADIOMETER ACCESSORIES

10.1 INTRODUCTION

The combination of a SAW coherent processor and four DVI's gives an analysis system able to display integrated spectra and cross-spectra on a CRT display. In this chapter several items of ancillary equipment are described briefly. The additional units are a pulsar timing unit (PTU), a direct memory access (DMA) computer interface and associated high speed data communication link, and a dual-channel polarization receiver. The interface and receiver are necessary to expand the basic system into one capable of analysing part of the spectrum emerging from a radio telescope and supplying the resultant integrated spectral data to an observatory computer. The PTU is a control device which allows synchronization of the DVI's to a pulsar frame signal.

10.2 PULSAR TIMING UNIT

The trigger period of the "mark 1" DVI control module is 4 μs so a new integrated spectrum is written to each DVI buffer every $N \times 4$ μs, where N is the number of integrations set on the DVI thumbwheel switches. From the point of view of a data acquisition computer the buffers are sampled with a period of $N \times 4$ μs, assuming the computer responds to every "new data available" pulse (figure 8-3). The sample period $T_s = N \times 4$ μs is therefore the time bin width of the system although, at 60% compressive receiver duty cycle, the smoothing on integrated spectral noise is consistent with an integration period of $0.6T_s$.

In the case of a pulsar signal the ratio P/T_s (P is the pulsar period) is not necessarily integral. Although it is not strictly necessary to have an integral number of samples per period, on-line data reduction techniques such as superposed epoch integration are usually more efficient if P/T_s is an integer.

When the DVI control module "pulsar" mode is selected, the PTU forces the DVI's to present a data acquisition computer with an integral number of summed spectra each pulsar period. This is done by preventing the control module being re-triggered for a short interval during the pulsar period, the period being defined as the time interval

between successive PTU "frame" or "delayed frame" pulses (figure 10-1). Normal DVI triggering is re-commenced after the next "delayed frame" pulse. When observation over a whole period is undertaken, the dead time at the end of the frame is less than T_s. Note that the "pulsar out" signal is a replica of the "new data available" signal.

A block diagram of the PTU is shown in figure 10-2. The sine wave output of an HF synthesizer passes through a signal conditioning stage and into a divide-by-10^6 counter chain, the output of which produces a "frame" pulse at the apparent period of the pulsar. The large division ratio ensures that the period is derived with sufficient precision given the frequency range and precision restrictions of typical synthesizers.

To provide a pulse phase control, the "Q" outputs of the last two counter stages are compared with a number set in the front panel "delay" thumbwheel switches. When coincidence is detected, a "delayed frame" pulse is produced. By varying the setting of the "delay" thumbwheel switches the "delayed frame" pulse can be moved in increments of P/100 relative to the "frame" pulse, allowing the phase of a pulsar signal to be adjusted within the processing window. Both the "frame" and "delayed frame" pulses are available at the PTU front panel but neither are used directly as DVI control signals.

The second PTU counter counts the number of "new data available" ("pulsar out") pulses sent from the DVI control module to the computer (figure 10-1). When the count reaches the number set in the PTU "samples" thumbwheel switches (M), the "pulsar in" signal to the DVI control module is activated (figure 8-14) and is not released until the next "delayed frame" pulse. The effect is to hold the 74LS390 4 μs divider in the control module reset, preventing the module and the SAW spectrometer being re-triggered.

A "go" signal is transmitted from the computer to the PTU via the computer interface and needs to be active if DVI triggering is to be permitted. The signal allows the computer to stop the entire spectrum analysis process, perhaps in order to write the results of an on-line analysis to a mass storage device. When "go" is re-activated the DVI control module begins operating on the first "delayed frame" pulse, so successive analyses remain synchronized.

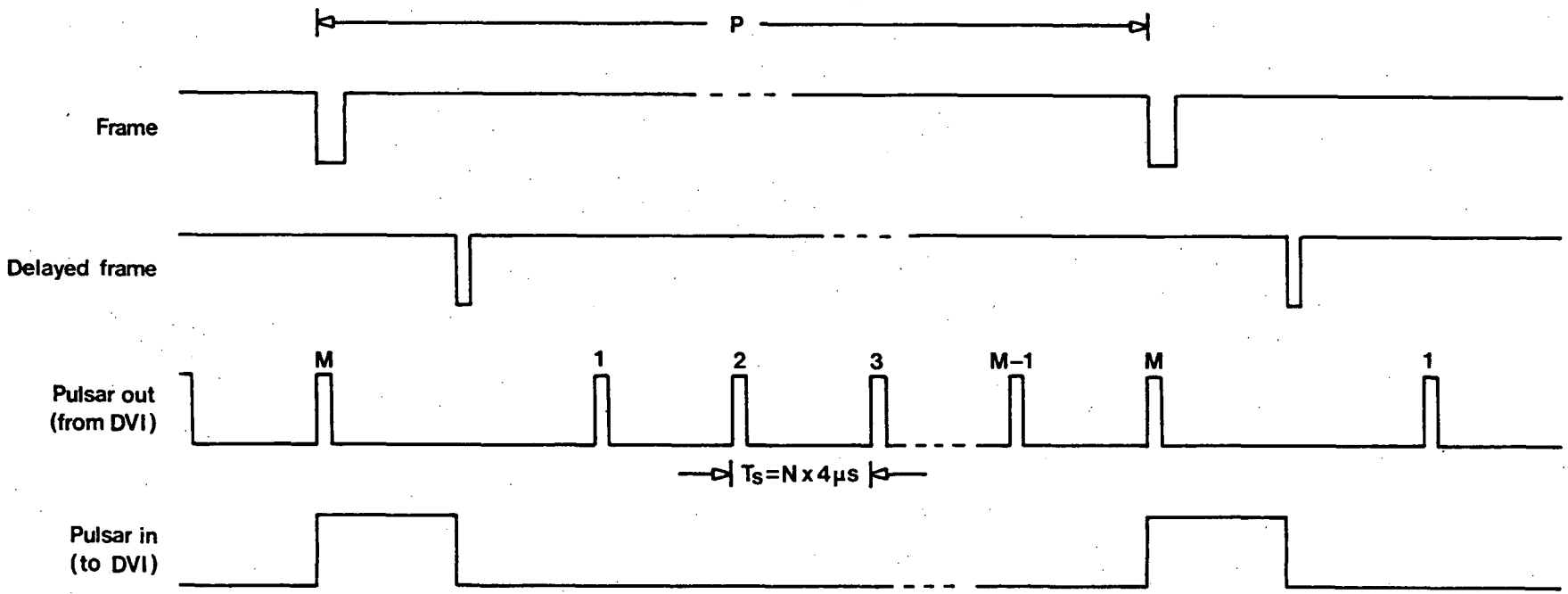


Fig.10-1: Timing diagram showing the operation of the PTU. When the number of "pulsar out" ("new data available") pulses matches the selected number of samples per period the "pulsar in" signal to the DVI control module is activated, preventing the module being re-triggered before the next "delayed frame" pulse.

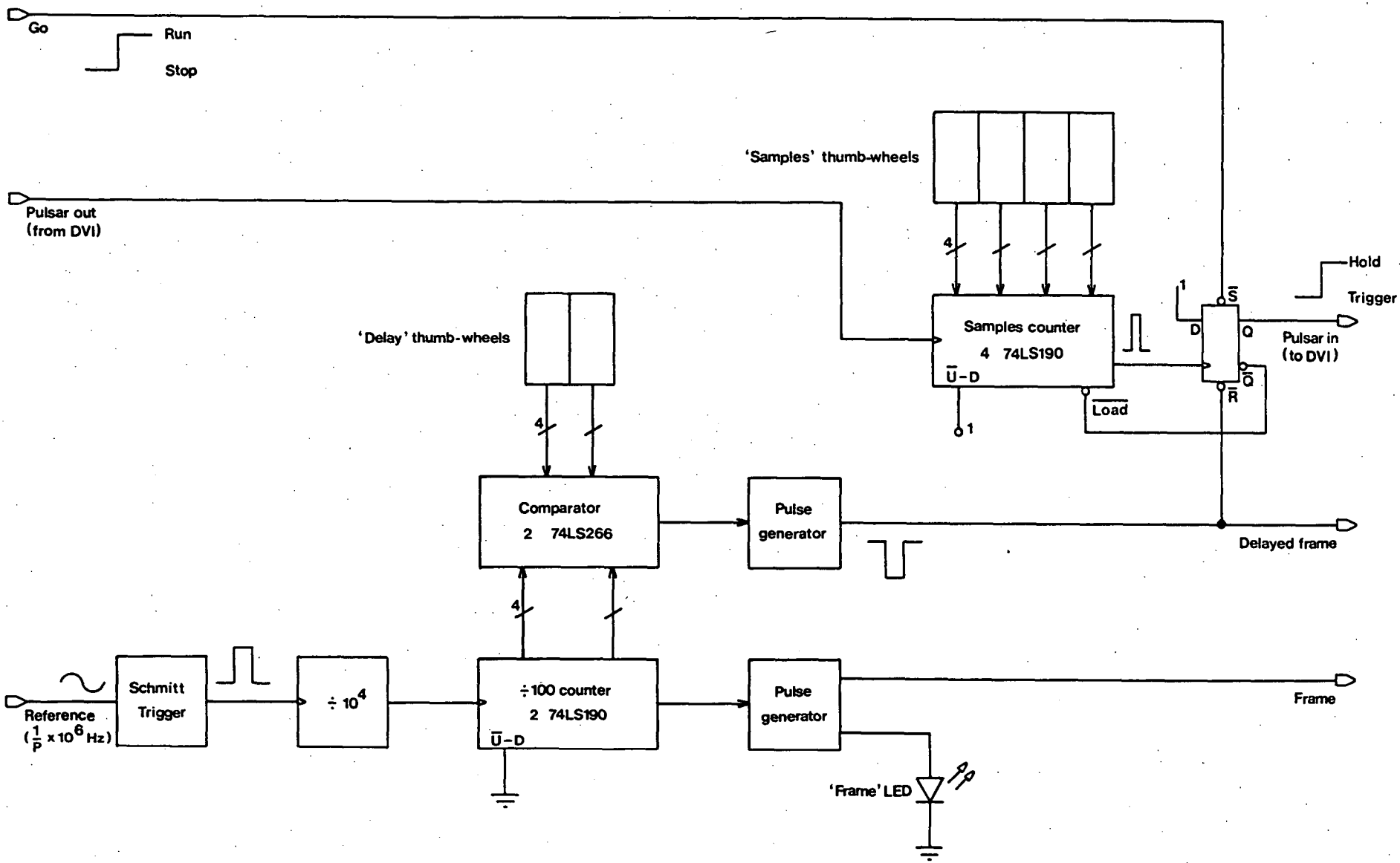


Fig.10-2: Block diagram of the PTU.

The PTU design permits the observing system to be used in a burst sampling mode if only a portion of the pulsar period needs to be examined. By adjusting the "delay" and "samples" controls, any time window of interest out of the period can be selected for sampling. For a given computer processing speed, burst sampling allows higher time resolution sampling than would be possible if the whole period were processed.

10.3 COMPUTER INTERFACE AND DATA COMMUNICATION LINK

The highest time resolution obtainable from a SAW radiometer when interfaced to a computer is equal to the time taken by the computer to acquire, process and store data. It is therefore advantageous to make these processes as efficient as possible. The four "mark 1" DVI output buffers, each of which is referred to as a "bank" of RAM, store a total of 256 sixteen-bit words. Disregarding computer processing or storage time, the time taken to acquire the contents of the four banks should not exceed a few hundred micro-seconds if the SAW coherent radiometer is to yield more than 100 samples per period on short-period pulsars such as PSR 0833-45.

The instruction execution times of typical mini-computers mean that a program controlled interface would yield a best time resolution of several milli-seconds rather than hundreds of micro-seconds. The alternative to a program controlled interface is a direct memory access (DMA) type. A DMA interface is more complex but allows the highest possible data rate, subject only to memory speed and bus timing restrictions within the computer.

In a DMA process the computer CPU is not involved in data acquisition operations once the transfer of data has been initiated. Blocks of data are sent directly into memory and functions normally undertaken by the CPU during memory operations are performed by interface hardware. At the end of an acquisition cycle the DMA interface returns control of the memory and bus to the CPU.

The DMA interface described in this section has been designed to allow up to four "mark 1" or "mark 2" DVI's to be connected to computers using the Digital Equipment Corporation (DEC) UNIBUS bus system.

Most PDP11 and VAX computers can host the unit.

The interface is in two parts which are connected by the data communication link. The "main" interface is a slightly-modified copy of a DEC DR11B DMA type and is located in or near the computer. The "device" interface is a new special-purpose PCB located in the enclosure containing the DVI's. Operation of a DR11B interface is described by DEC (1975) and since the device interface operation is associated closely with that of the main interface, a detailed description is not attempted in this section. Instead, the capabilities of the complete data transfer system are outlined briefly.

Data transmission takes place via an optically isolated, twisted pair, ribbon cable which is about 15m in length. Sixteen data bits and eight status bits can be transmitted to the computer. Eight control bits can be sent from the computer to the device interface and the PTU. Many of the status and control bits are read from, and written to, a register accessible to the observing program.

In practice, the status and control bits allow the two parts of the interface to communicate during the data transfer operation. As well, the status bits give the main interface access to some device interface internal signals, including several which are used in a data transmission error detection role. Information on the position of various DVI front-panel switches is also conveyed using the status bits. Apart from transmitting interface signals the control bits allow the observing program to read the setting of the DVI "Set N" thumbwheel switches, start or stop the DVI triggering via the PTU "go" signal, or illuminate an "N too small" warning LED on the DVI control panel. The LED is illuminated if the computer is unable to transfer the required data in the interval between successive "new data available" signals.

Data can be transferred in either of two program selectable modes: fast or slow DMA. In fast mode the interface has exclusive use of the UNIBUS during the transfer and data are sent as a single block (256 words for four banks). When slow mode is used the UNIBUS is shared with other DMA peripherals and DVI data are transmitted in several blocks, the size of which depend on the number of other peripherals requesting UNIBUS access. Fast mode is normally used in

single-user computer systems and typically gives a 20-30% advantage in transfer time. When data are transferred at full UNIBUS speed (400 ns/word) the best radiometer time resolution is about 150 μs, assuming a four-bank transfer. It is not always necessary to transfer data from all four DVI's and transmission from either 1, 2, 3 or 4 units can be software selected by the observing program.

10.4 POLARIZATION RECEIVER

A receiver designed for use in a coherent polarimetry system must amplify and frequency translate the outputs of two antennas sensitive to opposite polarization modes. Both channels in the receiving system need to be phase and amplitude matched as closely as possible. The closer the matching, the more likely it is that calibration information logged during an observing session will allow raw data to be corrected for the effects of residual mismatch.

It is not difficult to construct a narrowband, dual-channel receiver with accurate matching between the channels. In any case, the user of a narrowband polarimeter is free to insert attenuators and phase shifters into signal paths in an attempt to correct amplitude and phase errors. The amplitude and phase slopes of such devices across the narrow passband are usually not large enough to be a problem. Indeed, a common method of phase compensation is to use differential lengths of coaxial cable in the two receiver signal paths.

The problem of differential phase in a wideband receiver is illustrated in figure 10-3. It is assumed that the phase error characteristic between the two channels is linear but, because of the group delay responses of devices within the system, cannot be extrapolated to zero phase at zero frequency. Differential cable lengths produce a family of lines in the phase-frequency plane but the slope of the line giving the correct centre frequency phase shift does not match the slope of the error characteristic. If the phase error is delay line compensated at centre frequency, the passband edge error depends on the receiver bandwidth.

With a linear phase error it is clear that there is a delay line length which, when combined with a constant phase shift, produces the required phase-frequency compensation characteristic. A constant

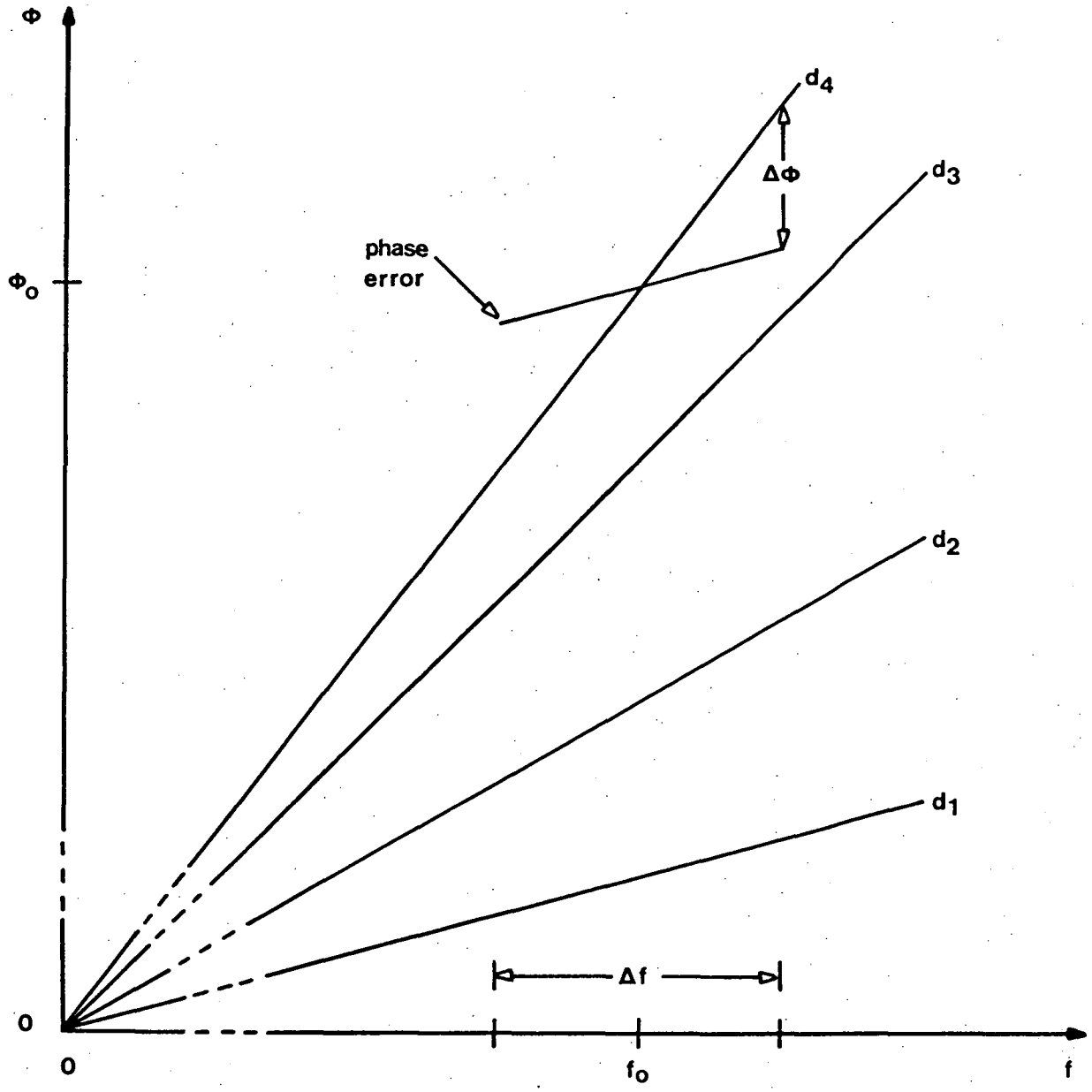


Fig.10-3: Phase lag characteristics for delay lines of various lengths. If a wideband dual-channel receiver exhibits a linear differential phase error characteristic such that the phase offset at f_o is ϕ_o, the delay line d_4 provides compensation at f_o with a band-edge error of $\Delta\phi$. For the error response shown a better means of compensation would involve the use of the delay line d_1 and a constant phase shift.

phase shift can be produced by inserting a phase shift element into a receiver CW local oscillator output so, in practice, a linear differential phase characteristic is permissible in a wideband receiver.

Other than in an LO phase adjustment role, most lumped-element variable phase shifters (including electronically or mechanically variable reflection types) are not suitable for use in wideband receivers. The passband group delay characteristic of these networks is normally not constant and the amplitude response is usually not uniform. The group delay is invariably a strong function of the mid-band phase shift, giving the units a large effective length range in transmission line terms. Even if identical phase shifters are placed in both receiver channels to compensate partially for the insertion phase of one device, it is difficult to manipulate the controls of three units (two signal, one LO) to give reasonable response matching.

A line stretcher is an acceptable means of introducing a variable phase shift into a signal path. The passband amplitude slope of such devices is very small and is usually insignificant even in wideband receivers. Depending on the slope of the linear differential phase response to be corrected, a second phase shifter may be needed in an LO line to give good band-edge compensation.

Figure 10-4 is a simplified diagram of a polarization receiver constructed for use with the SAW processor described in chapter seven. The centre frequency of 680 MHz coincides with the optimum working frequency of the Parkes observatory horn feed. Broadband RF design techniques are used throughout the receiver, so the operating frequency is determined mainly by the front-end filters. A single conversion design was adopted because the front-end filter rejection guarantees high attenuation (> 50 dB) at the 560 MHz image band.

Extensive use is made of TRW CA2820 broadband amplifier modules (section 7.3) and precision attenuators are incorporated in each IF channel. Low noise Avantek AK1000M amplifiers follow GaAs FET front-end amplifiers (constructed by P.M. McCulloch), giving the receiver an equivalent noise temperature of about 45 K un-cooled and 20 K when cooled to liquid nitrogen temperature. The two front-end filters are five-pole interdigital types, carefully matched in amplitude and phase responses.

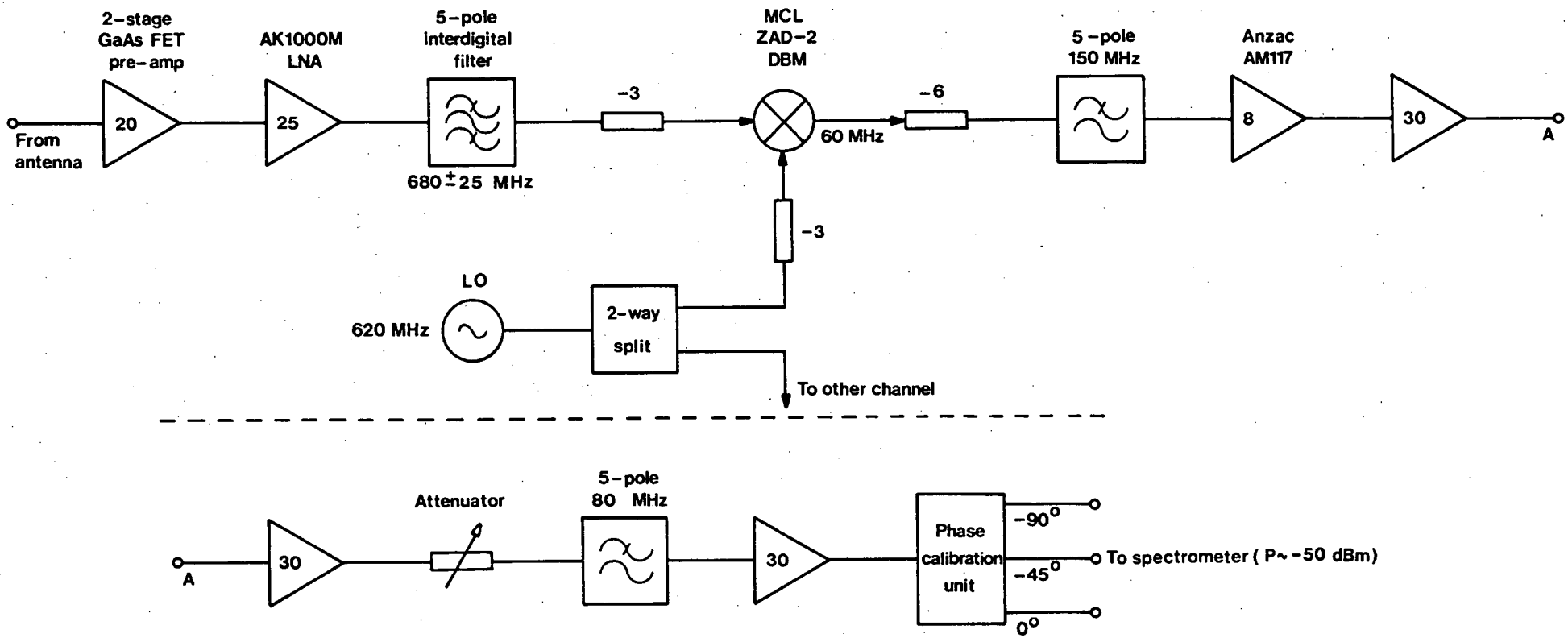


Fig.10-4: One channel of the wideband polarization receiver designed for use at the Parkes observatory. Components drawn above the broken line are mounted in the aerial cabin; those below the line are in the main receiver chassis located in the control room. Phase adjustment is by means of a line stretcher in the LO path to one channel and a delay line in one signal path. All 30 dB amplifiers are TRW CA2820 types.

Precision phase calibration networks at the two IF outputs are included to allow easy checking of the radiometer phase alignment during observations. Each unit produces a nominally frequency independent phase shift of 0^{o}, -45^{o} or -90^{o} and two identical networks are used to compensate as nearly as possible for the insertion phase and amplitude properties of a single unit. Figure 10-5 is a diagram of the broadband phase shifters. Figure 10-6 shows the amplitude responses obtained from one network and the worst-case difference response between like ports of the two units. In figure 10-7 deviations around the nominal phase values are shown for one calibration network. The worst-case differential response is also plotted.

The overall matching of the two receiver channels is very good, the phase error being plotted in figure 10-8. Measured from the Avantek amplifier inputs, the differential phase does not exceed $\pm 2^{o}$ over a 40 MHz bandwidth. Over the same range the differential amplitude is within ± 0.25 dB. Great care was taken with filter construction, filter response matching, correct device termination and matching of connecting lead lengths, but no phase or amplitude compensation was needed to obtain this performance.

The GaAs FET pre-amplifiers are normally aligned to give optimum noise figures, response errors between the two units then being compensated for in later stages. Measurements show that there is no difficulty in obtaining a substantially flat differential amplitude response and a linear differential phase response over a 50 MHz bandwidth, allowing effective compensation. Practical alignment and calibration procedures for a wideband polarimetry system are discussed in the next chapter.

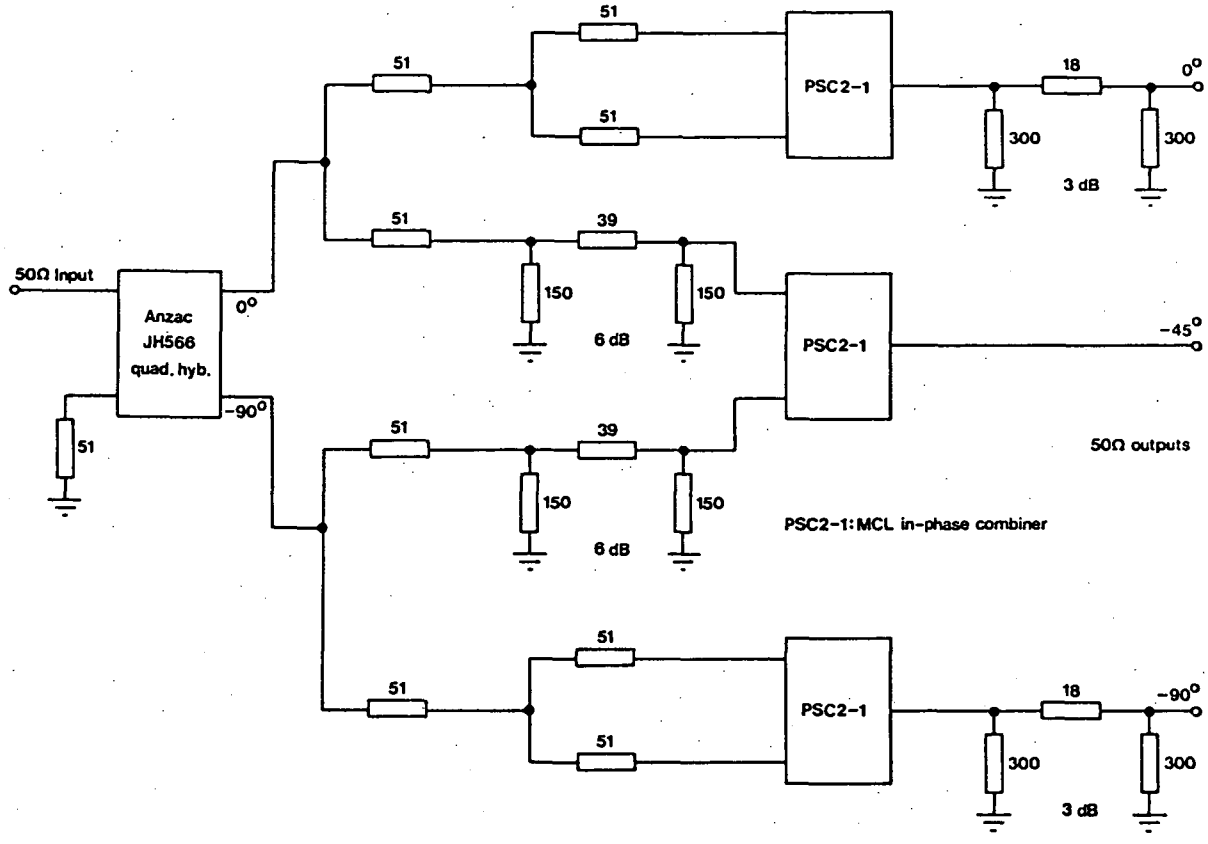


Fig.10-5: Circuit diagram of the IF phase calibration unit. Standard-value resistors are used.

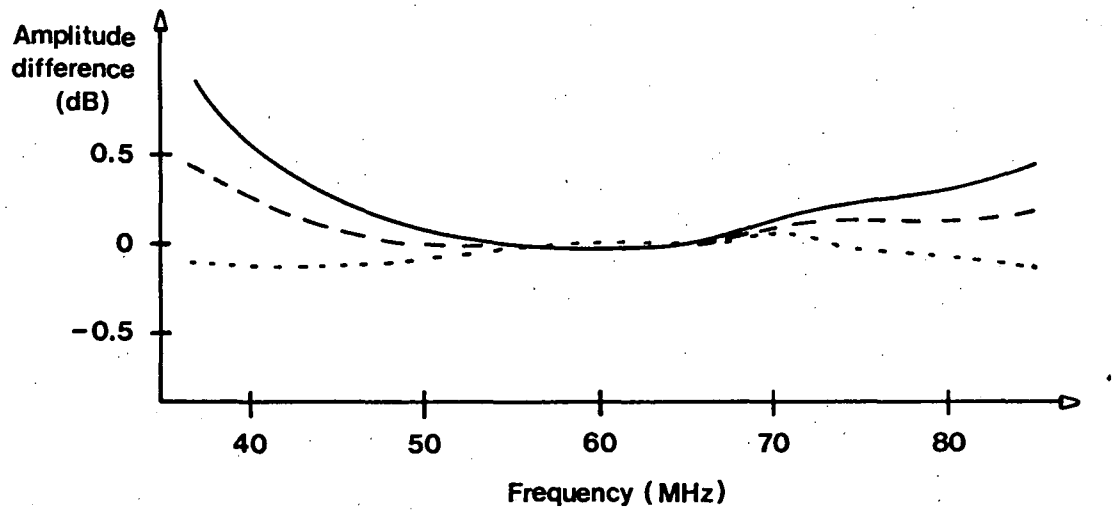


Fig.10-6: Differential amplitude response curves for the circuit of figure 10-5. The solid line represents the difference between the 0° and -90° ports; the dashed line is the difference between the 0° and -45° ports. The dotted line shows the worst-case difference between like ports in the two phase calibration units.

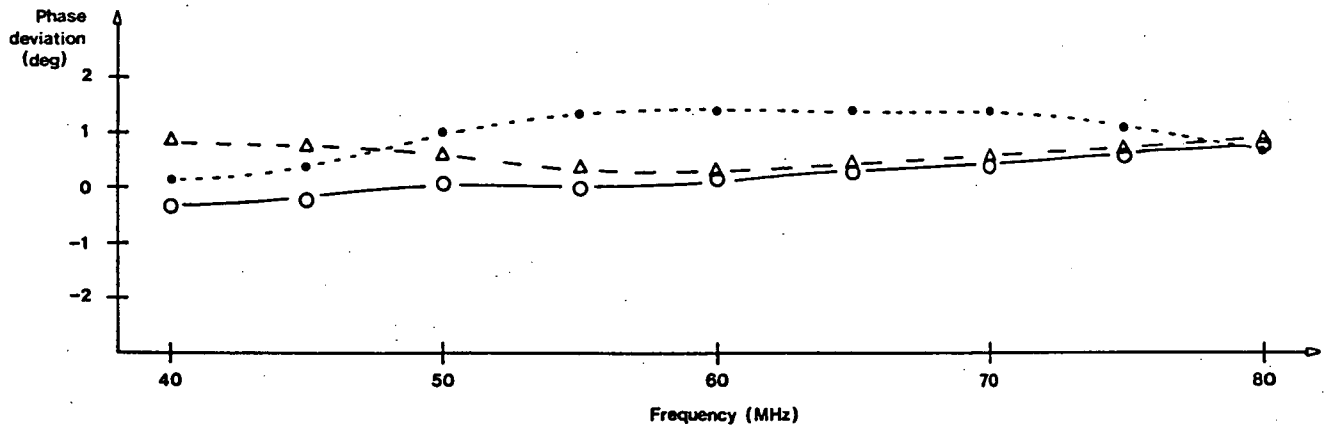


Fig.10-7: Phase error responses of the circuit shown in figure 10-5. The circles and triangles show the 0° to -90° and 0° to -45° responses respectively. The dots represent the worst-case error response between like ports in the two calibration units. Phase measurements in this and the next figure were made using an HP 8405A vector voltmeter (± 1.5° phase accuracy).

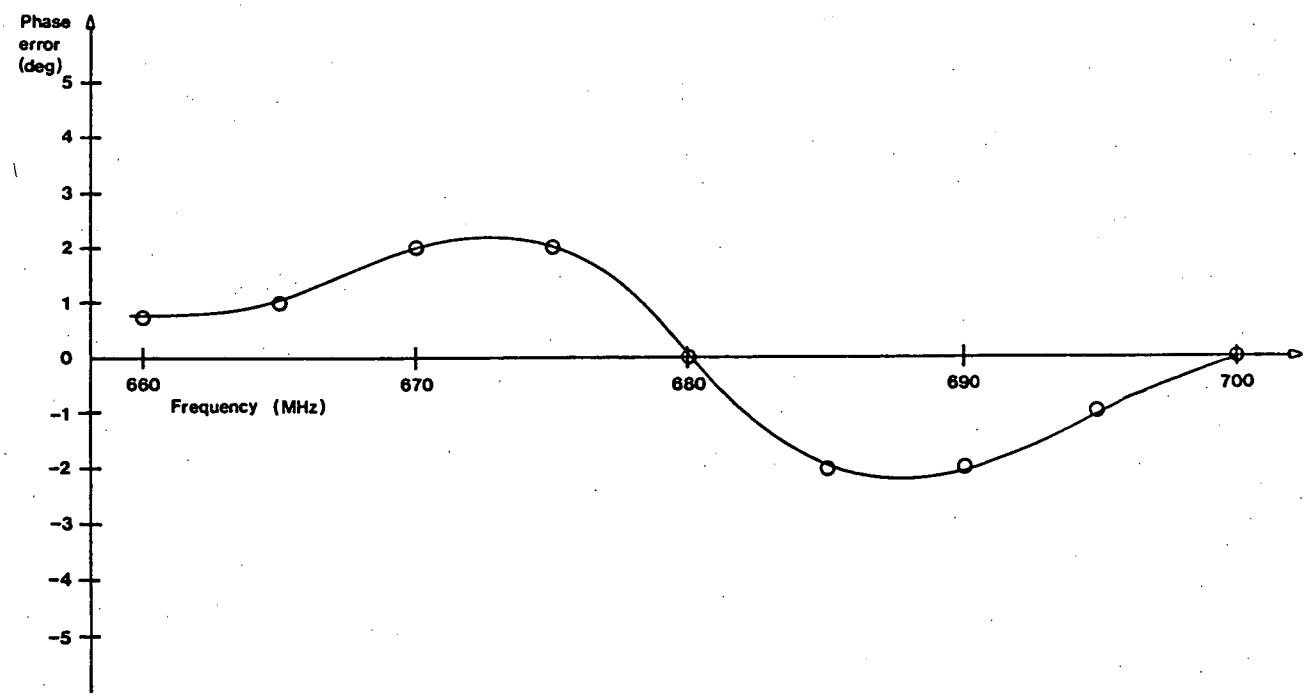


Fig.10-8: Differential phase response of the polarization receiver shown in figure 10-4. The characteristic was measured between the 0° ports of the IF phase calibration units. The in-phase test signal was applied to the inputs of the two Avantek amplifiers.

CHAPTER 10 - REFERENCES

Digital Equipment Corporation: "PDP11 peripherals handbook". 1975 ed., pp.7.13-7.21.

CHAPTER 11 INITIAL ASTRONOMICAL DATA

11.1 INTRODUCTION

During development of the equipment described in this thesis some astronomical data were collected using the Parkes radio telescope. These data were useful in verifying correct operation of the basic system but a number of problems involving both the new instrumentation and existing observatory hardware prevented scientifically useful results being obtained. This chapter describes the first scientific investigation conducted using the equipment in its final form. Observations were made with the 14 m steerable paraboloid at the Llanherne (Hobart) radio observatory. The aims of the investigation were:

(a) to verify the radiometer efficiency measurement reported in section 9.2 in an observing environment;
(b) to deduce current values for the dispersion measure (DM) and rotation measure (RM) of the Vela pulsar, PSR 0833-45.

Hamilton et al. (1977) showed that the RM of PSR 0833-45 was increasing at a rate of about 0.77 rad $m^{-2}yr^{-1}$ and speculated that the change was caused by a region of excess electron density containing a relatively strong magnetic field moving across the path to the pulsar. At the time of their 1976 study, evidence for a corresponding DM alteration was extremely tentative. The latest results show that there has been a significant DM change since 1970 and a further RM change since the 1976 study. Apart from early observations which demonstrated variations in the DM of the Crab pulsar of the order of 0.01% (Rankin and Counselman, 1973), the latest investigation gives the first conclusive evidence of a DM variation for any pulsar. The change amounts to more than 1% of the DM value deduced in 1970 by Komesaroff et al. (1971) and is thus two orders of magnitude greater than the variations reported for the Crab pulsar.

The Llanherne observations were made during October 1984. The Vela pulsar was tracked for approximately four hours during each observing session and integrated profile Stokes polarization data recorded on disk using the observatory PDP11/34 computer system.

Data were subsequently transferred to magnetic tape and analysed using a second PDP11/34 system with graphics and plotting facilities.

11.2 THE LLANHERNE OBSERVATORY

The major characteristics of the Llanherne telescope are summarized in table 11-1.

Table 11-1 Llanherne Telescope Characteristics

Latitude:	$42^{o}50'$ S.
Longitude:	$147^{o}30'$ E.
Dish Diameter:	14 m.
Mount:	Equatorial.
Feed:	629 ± 15 MHz disc type mounted at prime focus. Two opposite linear polarization channels are available. The feed is not rotatable.
Feed Isolation:	> 25 dB (free space).
Calibration Method:	Nominally equal amplitude, in-phase signal injected into both receiver channels via a third probe on the disc feed.

The original dual-channel receiver at the observatory was designed for PSR 0833-45 timing studies, most other pulsars being too weak to be observed satisfactorily with the small telescope. The receiver characteristics are listed in table 11-2.

Table 11-2 Llanherne Timing Receiver Specifications

Number of Channels:	2 intensity.
Centre Frequency:	635.0 MHz.
Pre-detection Bandpass Characteristic:	Third-order Butterworth.
-3 dB Bandwidth:	250 kHz.
Radiometer Bandwidth:	315 kHz.
Detector:	Square-law type based on an MC1496 multiplier and an LM301 post-detection amplifier.
Time Constant:	330 μs.
Equivalent Integration Time:	660 μs.

In the original receiving system the two RF channels were not phase or amplitude matched. Prior to the investigations described in this chapter, existing front-end components were aligned on-site to produce acceptable channel matching. The receiving system shown in figure 11-1 was then installed. A major requirement during alignment and installation was that the timing experiment not be disrupted. The new receiver uses much of the hardware described in section 10.4. A double conversion configuration has been adopted to suit the timing receiver requirements while at the same time working within the limitations of available frequency generation and translation equipment.

11.3 POLARIMETER CALIBRATION

Calibration methods for the polarimetry system have been developed from techniques used to calibrate narrowband polarimeters and ordinary spectrometers. The only slightly unusual feature of the calibration process is the need to quantify frequency-dependent phase errors in the instrumentation. If a broadband signal having known spectral and phase characteristics is injected into the entire receiving chain the required gain and phase correction factors can be derived.

The simplest calibration procedure involves injection of a noise signal into the telescope feed antenna. In disc or horn antennas sensitive to orthogonal linear polarization modes, a signal from a calibration probe with an azimuthal offset of 45^{o} relative to the feed probes produces equal amplitude, in-phase outputs from the main probes. The power (and amplitude) responses of the two RF channels can be found by pulsing the noise source and examining the relative sensitivity in each spectrometer frequency cell. Gain correction factors can be derived immediately from the relative sensitivity measurements. The pulsed calibration signal can be treated as a pulsar, allowing the signal-to-noise ratio (SNR) of the calibrating signal to be improved by synchronous integration.

The noise source is an in-phase signal so in an ideal polarimeter the U correlator output is a maximum and the V correlator output is nulled. In practice, phase compensation involves adjustment

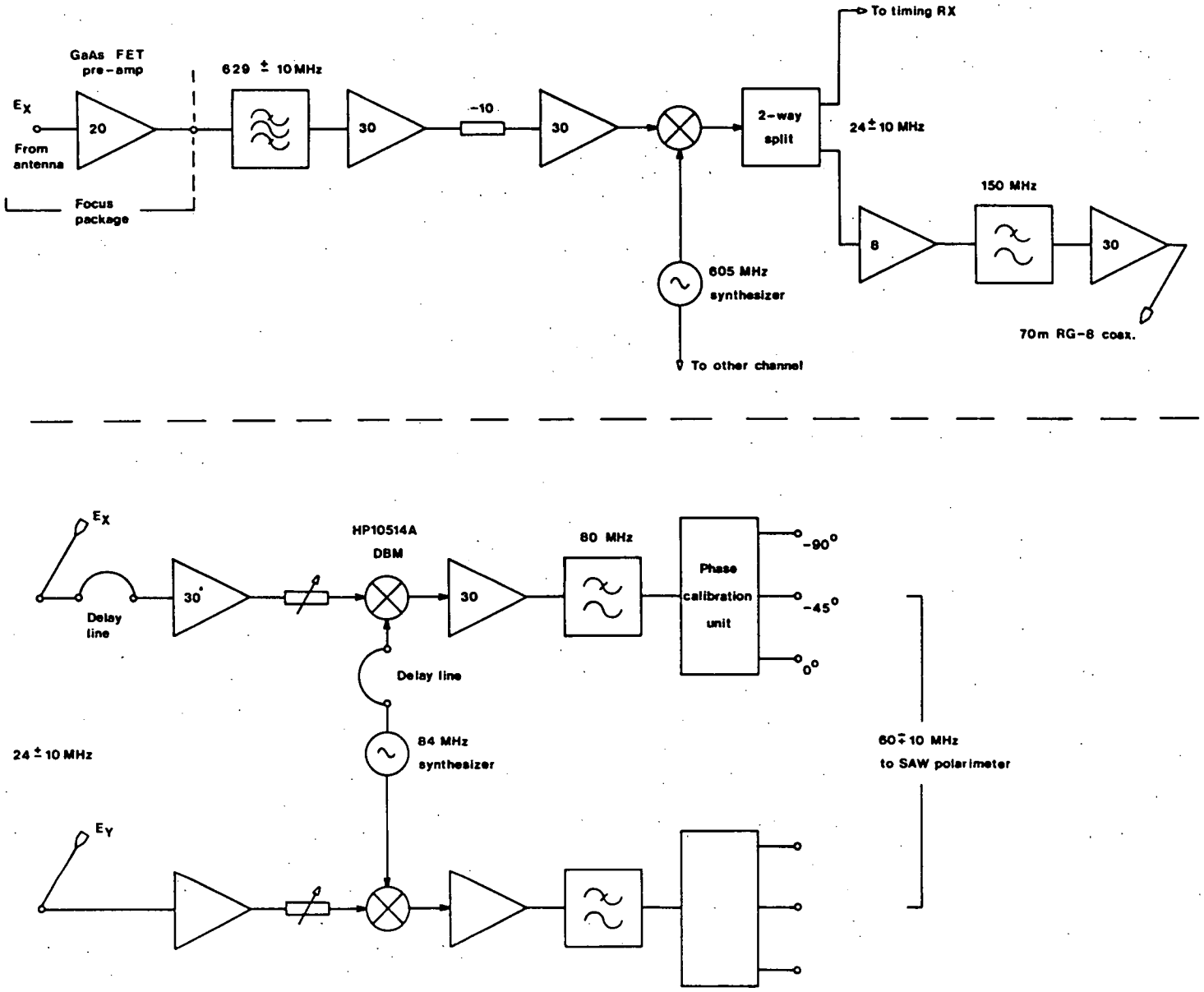


Fig.11-1: Block diagram of the Llanherne polarization receiver. Components drawn above the line are located in the 14m dish hut; those below the line are in the VHF array hut. The phase alignment is adjusted by changing the length of the delay lines.

of receiver phase shift elements until these conditions are approached as nearly as possible across the entire passband of the instrument. Insertion of 90^{o} and 45^{o} phase shifts (using the phase calibration unit described in section 10.4) then allows the gain ratio of the two correlators (nominally 1.0) to be deduced.

After the polarimetry system has been phase compensated, working calibration data can be obtained by using the U and V correlators as quadrature phase discriminators in a manner similar to that described in section 9.2. The residual phase error in the *ith* frequency cell is

$$\phi_i = \arctan\left(\frac{V_i}{U_i}\right) = \arctan\left(R_i\right) \quad , \qquad 11.3.1$$

where V_i and U_i are mean correlator outputs after correction for any correlator gain difference. The variance of the phase determination is:

$$\sigma^2_{\phi i} = \frac{R_i^2}{[1+R_i^2]^2}\left\{\frac{\sigma_{V_i}^2}{V_i^2} + \frac{\sigma_{U_i}^2}{U_i^2}\right\} \quad . \qquad 11.3.2$$

The power and differential phase responses of the complete Llanherne polarimetry system are shown in figures 11-2 and 11-3. The uncertainty in the power curves is less than 0.1 dB. Error bars on the phase plot are $\pm 2\sigma_{\phi_i}$ and were calculated using equation 11.3.2. Ripples in the characteristics are almost entirely due to the front-end filters. The passband attenuation slope of the RG-8 coaxial line (figure 11-1) is evident. Note that lower channel numbers correspond to higher front-end and first IF frequencies.

After gain corrections have been performed on data in a given cell, the residual phase error still produces distorted product Stokes parameters having the form

$$U'_i = 2\,\langle E_{1i}E_{2i}\cos\left(\delta_i + \phi_i\right)\rangle$$

and $\qquad 11.3.3$

$$V'_i = 2\,\langle E_{1i}E_{2i}\sin\left(\delta_i + \phi_i\right)\rangle \quad .$$

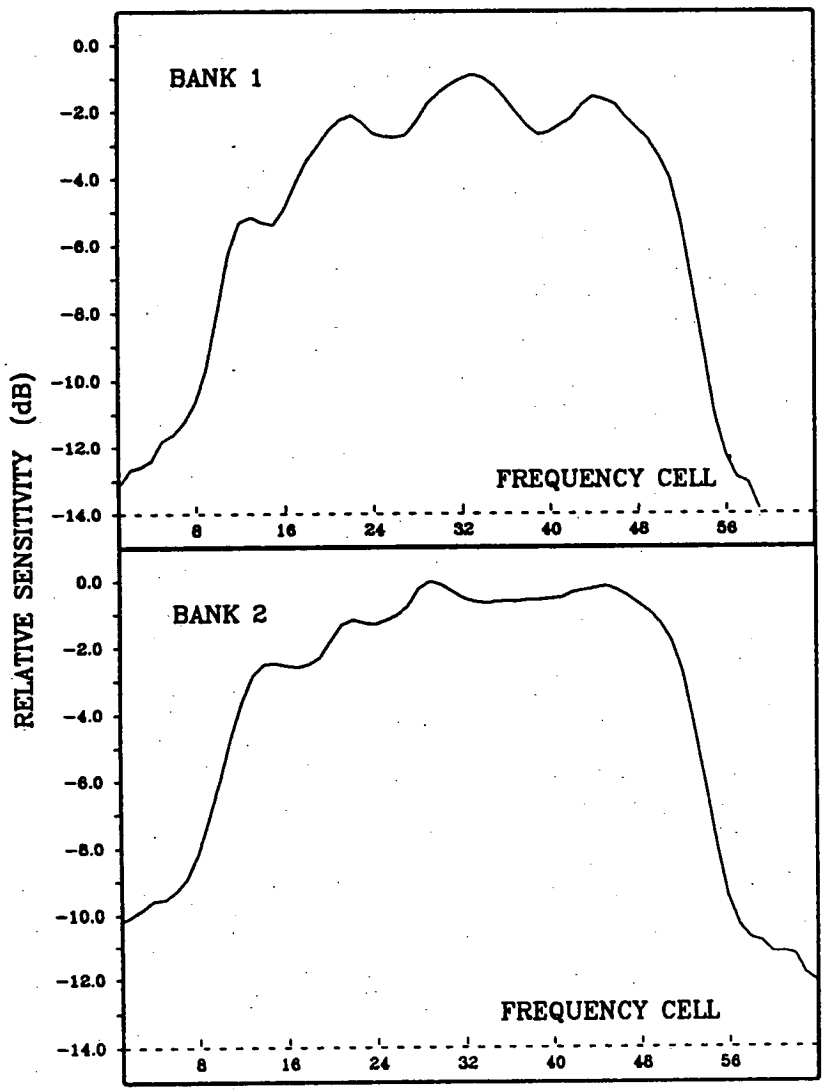


Fig.11-2: Responses of the two intensity channels in the Llanherne polarimetry system. The ordinates are referenced to the highest relative sensitivity value recorded.

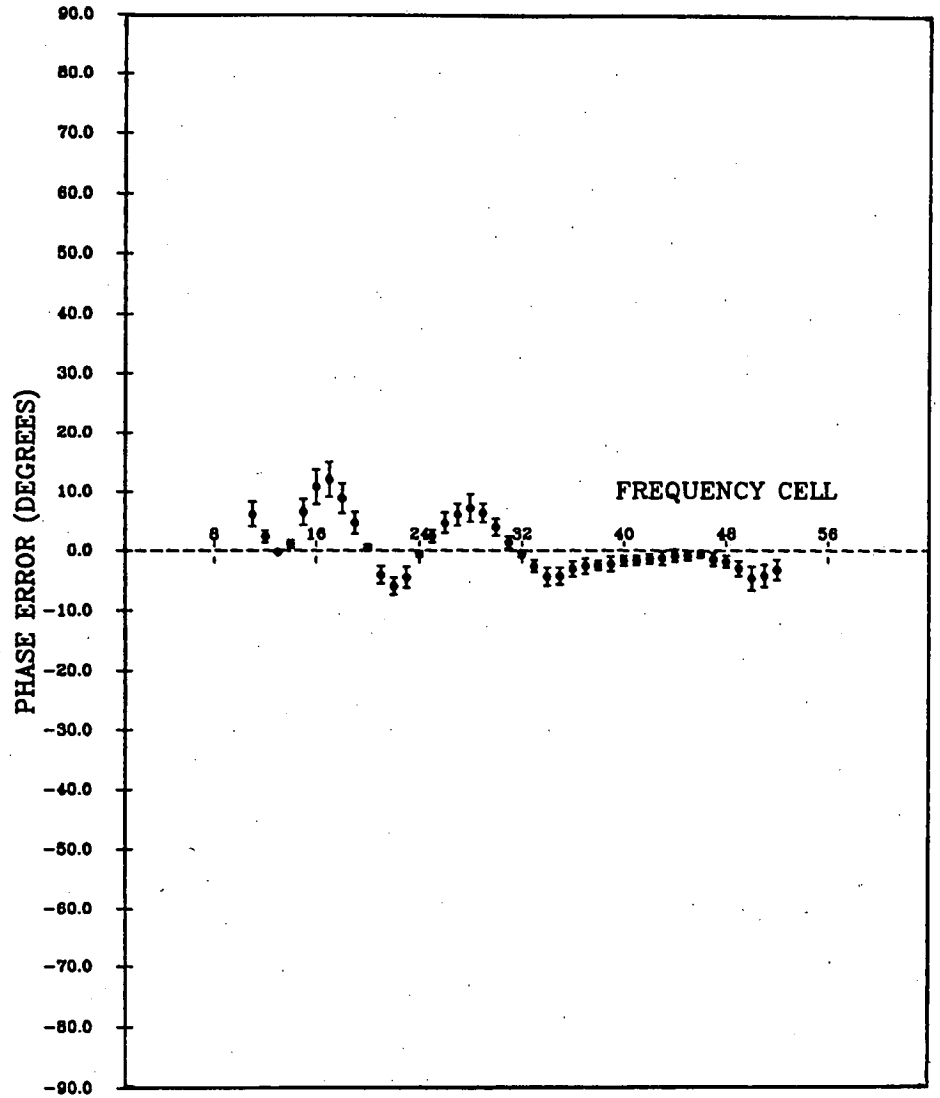


Fig.11-3: Differential phase response of the Llanherne system after phase compensation. Error bars are ± 2σ.

In these expressions E_{1i} and E_{2i} are the amplitudes of the orthogonal components in the i*th* cell and δ_i is their phase difference. Expanding the compound angles shows that the phase error results in a U'-V' measurement plane rotated away from the true U-V plane in the Poincaré representation (figure 1-1). Correct parameters are derived by rotating the U'-V' plane into coincidence with the U-V plane using the transformation:

$$\begin{pmatrix} U_i \\ V_i \end{pmatrix} = \begin{pmatrix} \cos \phi_i & \sin \phi_i \\ -\sin \phi_i & \cos \phi_i \end{pmatrix} \begin{pmatrix} U'_i \\ V'_i \end{pmatrix} \quad . \qquad 11.3.4$$

Calibration information was logged frequently during the Llanherne observing sessions. In addition to a pulsed noise signal, CW signals were often injected into the receiving system via the feed calibration probe. The CW signals allowed accurate determination of the observing frequency and were generated by a synthesized signal generator. To a measurement resolution of 1 part in 10^6 the synthesizer master oscillator was identical in frequency to a 1 MHz output from the observatory rubidium frequency standard.

An absolute flux scale was established for the Llanherne observations by referencing the intensity of the noise calibration signal to that of the unpolarized continuum source Virgo A (M87). A source flux of 400 Jy was assumed over the whole passband. The equivalent flux of the calibration signal was 360 Jy and it is expected that fluxes derived for PSR 0833-45 are accurate to within ± 10%.

11.4 SAW RADIOMETER EFFICIENCY AND PERFORMANCE

The efficiency of the SAW radiometer was assessed by comparing the total intensity SNR obtained when observing PSR 0833-45 with the SNR obtained from a simultaneous observation using the timing receiver. In both systems 1000 pulses were integrated, the results of the integrations being displayed on separate CRT displays.

Considering a single pulse and using the parameters given in table 11-2, the timing receiver bandwidth-time product is

$$\Delta f \Delta t = 315 \times 10^3 \times 660 \times 10^{-6} = 208.$$

At 60% duty cycle the post-detection integration time for the SAW radiometer is

$$\Delta t_{SAW} = 2.4N \times 10^{-6} \text{ s},$$

where N is the number of integrations set on the DVI thumbwheel switches. The SAW analyser radiometer bandwidth (Δf_r) is 950 kHz (section 9.2) so

$$\Delta f_r \Delta t_{SAW} = 950 \times 10^3 \times 2.4N \times 10^{-6} = 2.28 \text{ N}.$$

For equal bandwidth-time products,

$$\Delta f \Delta t = \Delta f_r \Delta t_{SAW} ,$$

giving N = 91.

It was clear from the CRT displays that the SNR in the two systems was nearly the same for N = 91. Subsequent analyses of 1000-pulse integrations taken over a period of hours showed that the relative sensitivity of the SAW instrument was 87 ± 2%. The comparison was made using data from spectrometer frequency cells around 635 MHz (the timing receiver frequency) but results were similar when data from other cells in the passband were used. Several integrations made using different values of N showed the expected $\sqrt{N}$ variation in sensitivity.

The result of the sensitivity comparison is better than expected given the laboratory measurement of 72% SAW radiometer efficiency (section 9.2). However, the timing receiver passband characteristics may not conform exactly to specification, causing the radiometer bandwidth to be different from the figure given in table 11-2. In addition, it is probable that the conventional receiver has imperfections of its own to be accounted for. The important point is that the SAW radiometer produces demonstrably acceptable results in an actual observatory environment.

As well as efficiency measurements, attempts were made to assess the instrumentation stability and suitability for pulsar

observations. The largest instability noted was a shift in the observing centre frequency of 10 kHz when the ambient temperature changed by more than 3^{o}C. This change amounts to 0.015 resolution cell or 2 parts in 10^{5} relative to the nominal observing frequency. The resolution in detecting frequency changes was approximately 1 kHz and, at this limit, no corresponding change in the output time-frequency scale of the SAW spectrometer occurred. Measurements of the DVI master oscillator frequency over a period of several weeks indicated no frequency change at a measurement resolution of 1 part in 10^{6}. Small gain and phase changes in the observing system were attributed to changes in conventional RF stages, particularly the GaAs FET pre-amplifiers. Any contributions from the SAW processor itself were masked by these changes. Finally, post-detection and post-correlation d.c. levels required no adjustment after a warm-up period of about one hour.

The overall performance of the prototype SAW coherent radiometer leaves no doubt as to its suitability for pulsar radioastronomy. In experiments where the exact observing frequency is important it is advisable to check the frequency each time noise calibrations are undertaken. This procedure was followed at Llanherne. There is considerable scope for further characterization of the instrumentation performance in other types of astronomical observation.

11.5 DISPERSION AND ROTATION MEASURES OF PSR 0833-45.

Data from which DM and RM determinations were made were collected over a four hour period around transit on 23 October 1984. The observation period corresponded to the interval 0500 - 0900 solar time. Data were collected in the form of 4000-pulse synchronous integrations, the results of each integration being written to a separate file. Individual integrations yielded a SNR in excess of 15. The sample interval was 670 μs and an accurate apparent pulsar period was derived by using a computer controlled HF frequency synthesizer as a PTU reference. Computer limitations restricted the number of spectrometer cells able to be observed simultaneously to 15 but different blocks of channels were observed in successive integrations.

Care was taken to ensure considerable overlap in frequency between data recorded in separate files. File concatenation procedures used during data processing could therefore be checked readily.

The process used to determine DM was basically one of producing a least-squares fit to the relative pulse arrival times in individual spectrometer cells. For two cells with centre frequencies f_1 and f_2 (MHz) the dispersion delay (equation 1.4.5) is

$$\Delta t_d = \frac{DM}{2.410\times10^{-4}} \left[\frac{1}{f_1^2} - \frac{1}{f_2^2}\right] \quad s \ .$$

Thus, a plot of Δt_d versus $X = \dfrac{\left[\dfrac{1}{f_1^2} - \dfrac{1}{f_2^2}\right]}{2.410\times10^{-4}}$ for fixed f_1 or f_2

is a straight line of slope DM. In the present study, the centroid of the highest-frequency cell (cell 1) was chosen as an arbitrary reference frequency. Linear regression analysis was then performed and the slope of the best-fit line taken as the best DM estimate.

Cross-correlation was used to determine the dispersion delay between a template profile and profiles in each cell. The template was formed by de-dispersing the total intensity profiles in each file, the catalog value of DM = 69.08 being used for the first de-dispersion. The total intensity profile in each cell was then cross-correlated with the template to give a relative arrival time. Linear regression then gave an amended DM. The new DM was used to de-disperse the profiles, regression performed and yet another DM obtained. This iterative process continued until the calculated DM became stable. The procedure produced rapid convergence even when very unrealistic DM values were tried as a starting point. The arrival time versus X characteristics from all files were then concatenated and a final characteristic produced. Standard errors in intermediate DM values were about ± 0.3 and direct averaging of intermediate values produced a DM result close to that obtained from the final delay plot.

Figure 11-4 shows a typical correlation template as well as de-rotated, de-dispersed polarization parameters. Correlation coefficients of more than 0.9 were always obtained in the arrival time

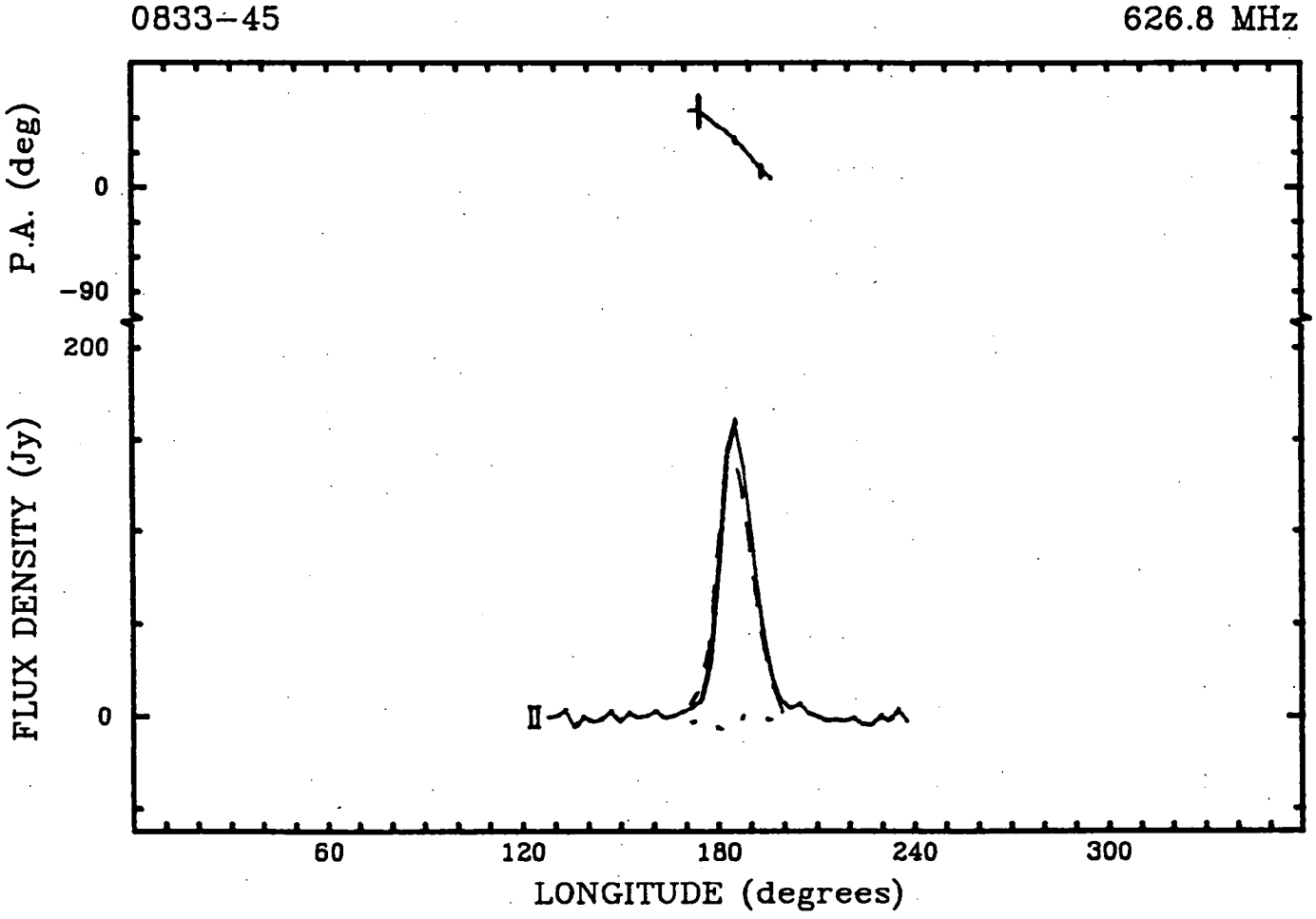


Fig.11-4: A typical de-rotated, de-dispersed integrated profile (4000 pulses) for the Vela pulsar recorded using the Llanherne polarimetry system. Data from 15 frequency cells around 626.8 MHz were used to produce the profile. The total intensity (I) is shown as a solid line, the linear component (L) is shown dashed and the circular component (V) is represented by the dotted line. The error box is 1 sample bin wide and ± $2\sigma_I$ high. Position angle error bars are ± 2σ. Total intensity profiles were used as cross-correlation templates in the DM analysis.

analysis. Figure 11-5 shows the final plot of relative arrival time versus X. There is an arbitrary constant time offset due to the arbitrary pulse phase of the correlation templates. The best estimate of DM is 68.2 ± 0.2 (2σ error). The linear regression was un-weighted because all correlation coefficients were nearly the same.

To determine the pulsar RM, the Faraday effect was exploited directly. The basic Faraday rotation relationship (equation 1.4.2) gives the rotation angle for linearly polarized radiation as

$$\theta = RM\lambda^2 \qquad \text{radian.}$$

A plot of the position angle (PA) of the pulsar radiation (at some fixed longitude) versus λ^2 therefore gives a straight line of slope RM. The value deduced is the total RM along the line of sight to the pulsar and includes a contribution from Earth's ionosphere.

Two modifications were made to the basic method. The first change was the manner in which the PA waveform offset between two frequency cells was determined. Rather than choosing a single fiducial longitude, the PA difference and the uncertainty in the difference were computed at several points across the pulse. A weighted average of these differences then gave an improved estimate of the PA offset. In practice, offsets were calculated relative to the highest-frequency cell in the usable portion of the passband.

The second modification involved compensation for ionospheric rotation. The ionospheric RM was calculated according to equation 1.4.4 and appropriate rotations applied to the U and Q Stokes parameters during data processing. The slope of the PA offset versus λ^2 graph therefore gives the pulsar RM directly. Ionospheric critical penetration frequency (f_0F_2) data were obtained from the Hobart office of the Ionospheric Prediction Service. The IPS data gave f_0F_2 values at 15 minute intervals over the period in which observations were made. Typical f_0F_2 readings were around 3 MHz, resulting in an ionospheric RM of about -0.6 at transit. Zenith angles were small (about 2^o at transit) so it is expected that ionospheric corrections were accurate to within ±5%. In any case, the corrections were always small relative to the catalog RM value of +33.6 for PSR 0833-45.

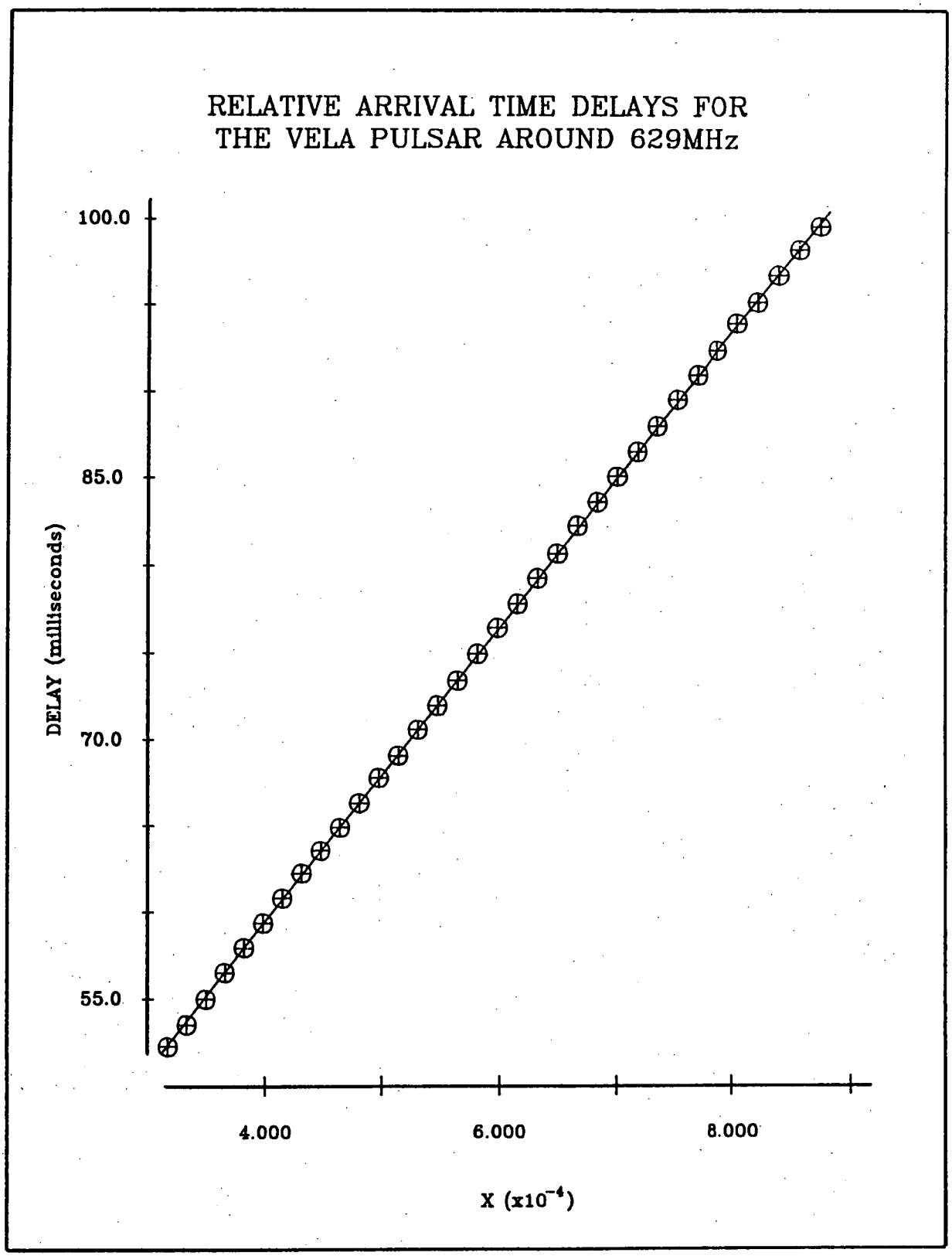


Fig.11-5: Relative pulse arrival time plotted as a function of $X = [(1/f_1^2) - (1/f_2^2)]/(2.410\times10^{-4})$ for PSR 0833-45. The line of best fit is shown.

Figure 11-6 is a plot of PA offset versus λ^2. Error bars are $\pm 1\sigma$ and weighted linear regression gives RM = 46.1 ± 1.6 (2σ error).

11.6 ASSESSMENT OF LLANHERNE DATA

The integrated polarization profile obtained in each polarimeter spectrometer frequency cell and the de-dispersed, de-rotated profile shown in figure 11-4 (and in more detail in figure 11-7) are similar to the 631 MHz profile (figure 11-8) published by McCulloch et al. (1978). In both observations a high degree of linear polarization (~80%) is present with the leading edge of the pulse being particularly highly polarized. The bandwidth used in the 1978 study was 180 kHz compared with 650 kHz (half-power bandwidth) in the present investigation. Thus, more dispersion smearing (~2%) is evident on the latest profile. The effect of smearing is also noticeable on the position angle waveform where changes in PA tend to be averaged out by dispersion. Note that position angles recorded at Llanherne are relative since the orientation of the disk feed on the sky was not determined.

A noticeable difference between the profiles is the lack of significant circular polarization in figure 11-7. It is probable that the Llanherne calibration probe does not produce exactly in-phase signals from the feed. If the polarimeter is aligned on the assumption of a correctly phased calibration signal, an error of about 4^o in the signal could result in cancellation of the small circularly polarized component from the pulsar. In essence, the U'-V' plane (section 11.3) is always rotated away from the U-V plane by a small angle. It is likely that, to the first order, the rotation is constant over the instrument passband. Cancellation could also be the result of cross-coupling in the feed. At 25 dB isolation the maximum value of $|V_{spurious}|/L$ is about 0.11 (section 1.3), sufficient to offset the pulsar circularly polarized component. In the absence of independent measurements of calibration signal phase or feed isolation (physical arrangements at Llanherne make in-situ measurements impractical), no adjustment has been made to data.

The ratio $L/|V|$ is high (~13) for PSR 0833-45 so the effect of small phase errors or of weak coupling on RM determination is expected to be slight. To verify this, deliberate spurious rotations

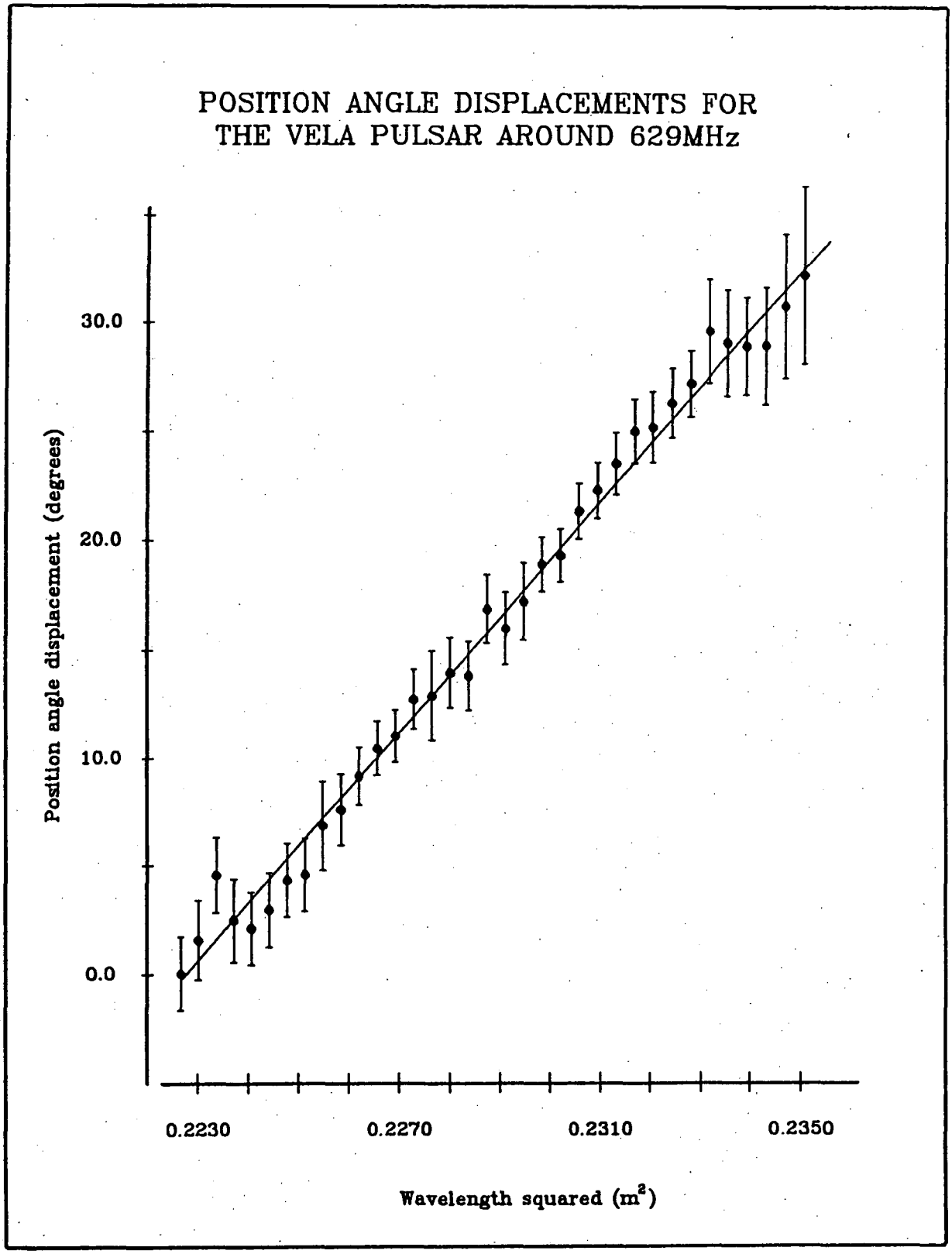


Fig.11-6: Relative position angle offset versus λ^2 for PSR 0833-45. Error bars are ± 1σ and the line of best fit is derived using weighted linear regression.

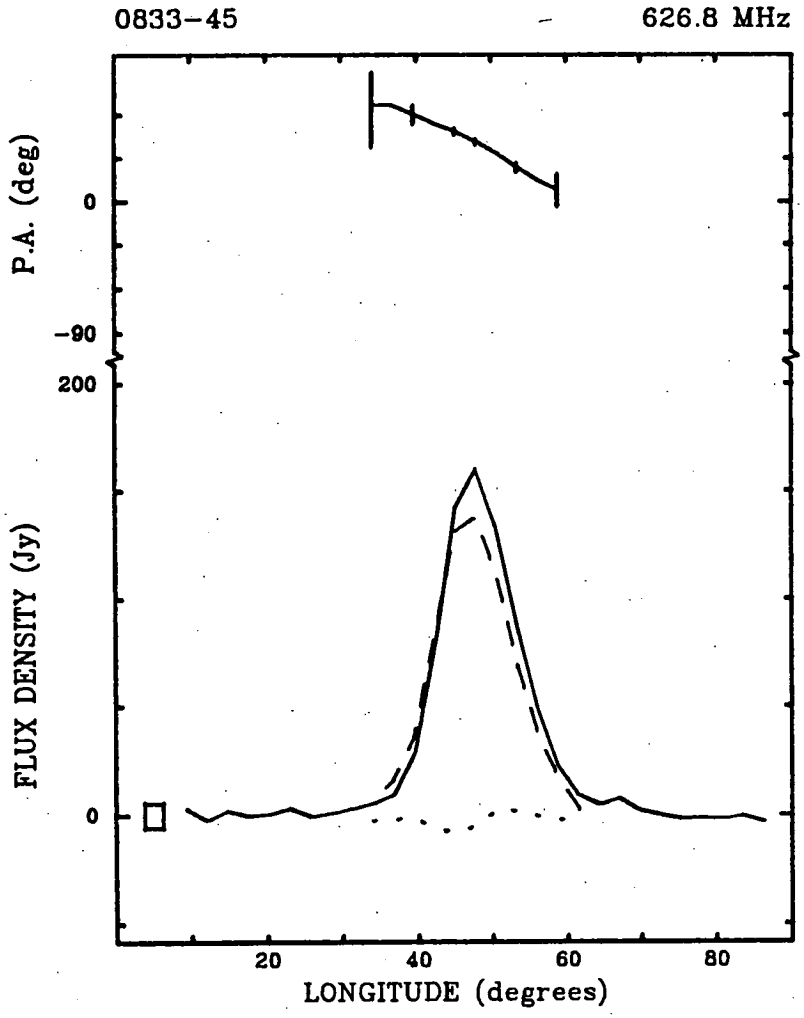


Fig.11-7: The de-rotated, de-dispersed profile of figure 11-4 shown in more detail. The sampling resolution is equivalent to approximately 130 samples per period.

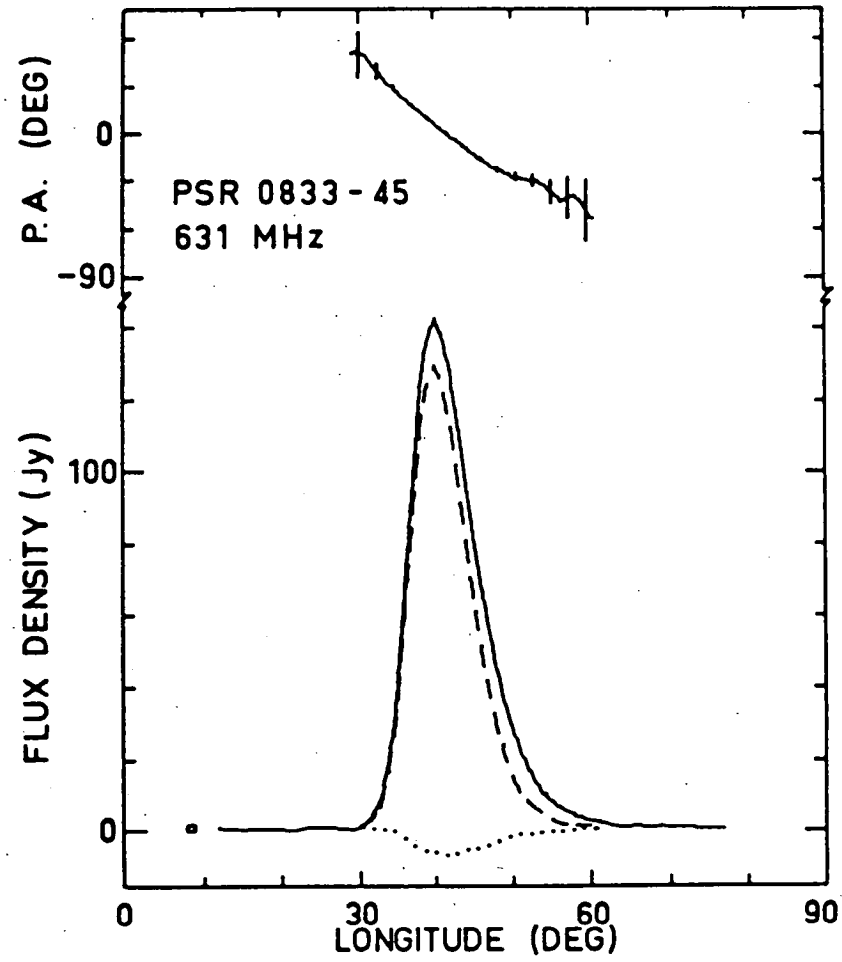


Fig.11-8: The 631 MHz profile (6000 pulses) for PSR 0833-45 recorded by McCulloch et al. (1978). Their observations were made using the 64 m Parkes telescope. The observing bandwidth was 180 kHz and the post-detection time constant was 200 μs.

of the U-V plane were introduced during processing. A rotation of 10^{o} (more than double the supposed error) produced RM changes of less than 0.5, within the RM error limits given earlier.

Phase errors have no effect on DM determination since total intensity profiles are cross-correlated during processing. To test the sensitivity of the entire DM measurement process to observing frequency errors, frequency offsets were applied during data processing. For errors of 10 kHz the resultant DM change was about 0.02, an order of magnitude less than the quoted DM error.

The sweep rate of PSR 0833-45 at 629 MHz is approximately 440 MHz/s so the pulsar instantaneous frequency moves about 1 kHz or 0.0015 resolution cell during the spectral sampling interval T_1 = 2.4 μs. For practical purposes the sampling is instantaneous and the SAW system is functionally the same as a filterbank. However, the SAW radiometer has the advantage that integrated spectra obtained from the DVI's apply to a given instant, regardless of the data acquisition speed of the computer. Thus, no correction is necessary for spurious dispersion effects such as the one introduced when a multiplexed A/D converter is used to sample filterbank channels sequentially.

Finally, the effect of interstellar scattering on the DM measurement needs to be considered. Scattering is mentioned again in the next section but at this stage it is noted that the pulse from the Vela pulsar is severely broadened by scattering at low frequencies and that the broadening follows the usual f^{-4} law. If the pulse phase at a given frequency is deduced from the position of the pulse peak then, to the first order, the excess phase due to broadening also follows an f^{-4} law. If scattering were significant, use of a cross-correlation technique to determine pulse phase would yield an artificially high value of DM, not the decrease observed.

Using the results of Ables et al, (1970), it is estimated that the extra "delay" due to scattering across the passband of the SAW polarimeter is less than 50 μs for PSR 0833-45. This is about 30 times less than the change in pulse delay which occurs if the DM changes from 69.08 to 68.2. Errors due to scattering are therefore insignificant in the present DM determination.

11.7 DISCUSSION

Figure 11-9 is a plot showing measured values for the DM and RM of PSR 0833-45 over a 15 year period. Data to 1976 are referenced by Hamilton, McCulloch, Manchester, Ables and Komesaroff (1977), hereafter HMMAK. August 1979 data are from an unpublished study by P.A. Hamilton. Changes in both DM and RM are evident and the general trends are shown by the straight lines which represent weighted least-squares fits to the data. The slopes of the DM and RM lines are -0.040 ± 0.004 pc cm^{-3} yr^{-1} and 0.73 ± 0.10 rad m^{-2} yr^{-1} respectively (2σ errors).

HMMAK cite evidence suggesting that the cause of the variations lies in the interstellar medium rather than in the pulsar magnetosphere. The implication of the DM and RM changes is then that a region of excess electron density has moved (and is probably still moving) out of the line of sight to the pulsar. Motion of the pulsar, the region of excess density, or both, could account for the changes. The RM is increasing so the magnetic field in the region must be directed towards the pulsar, giving a negative RM contribution. This means that the internal field is in a direction opposite to the mean interstellar field in the direction of PSR 0833-45.

HMMAK note that variation in DM and RM over a short time-scale suggests that the region of excess density must be quite small. They suggest that the Vela pulsar may have a transverse velocity as large as 500 km s^{-1}. Adopting this figure as a reasonable upper limit, the transverse scale length is about 5.1×10^{-4} pc or 105 AU per year. If the line of sight dimension is comparable, the average change in DM of 0.04 $cm^{-3}yr^{-1}$ implies an excess electron density of 78 cm^{-3}. This electron density is much higher than the average interstellar value of about 0.03 cm^{-3}.

The weighted mean line of sight component of the magnetic field in the region of excess density can be estimated as

$$\langle B_{\parallel} \rangle = 1.23 \frac{\frac{d(RM)}{dt}}{\frac{d(DM)}{dt}} ,$$

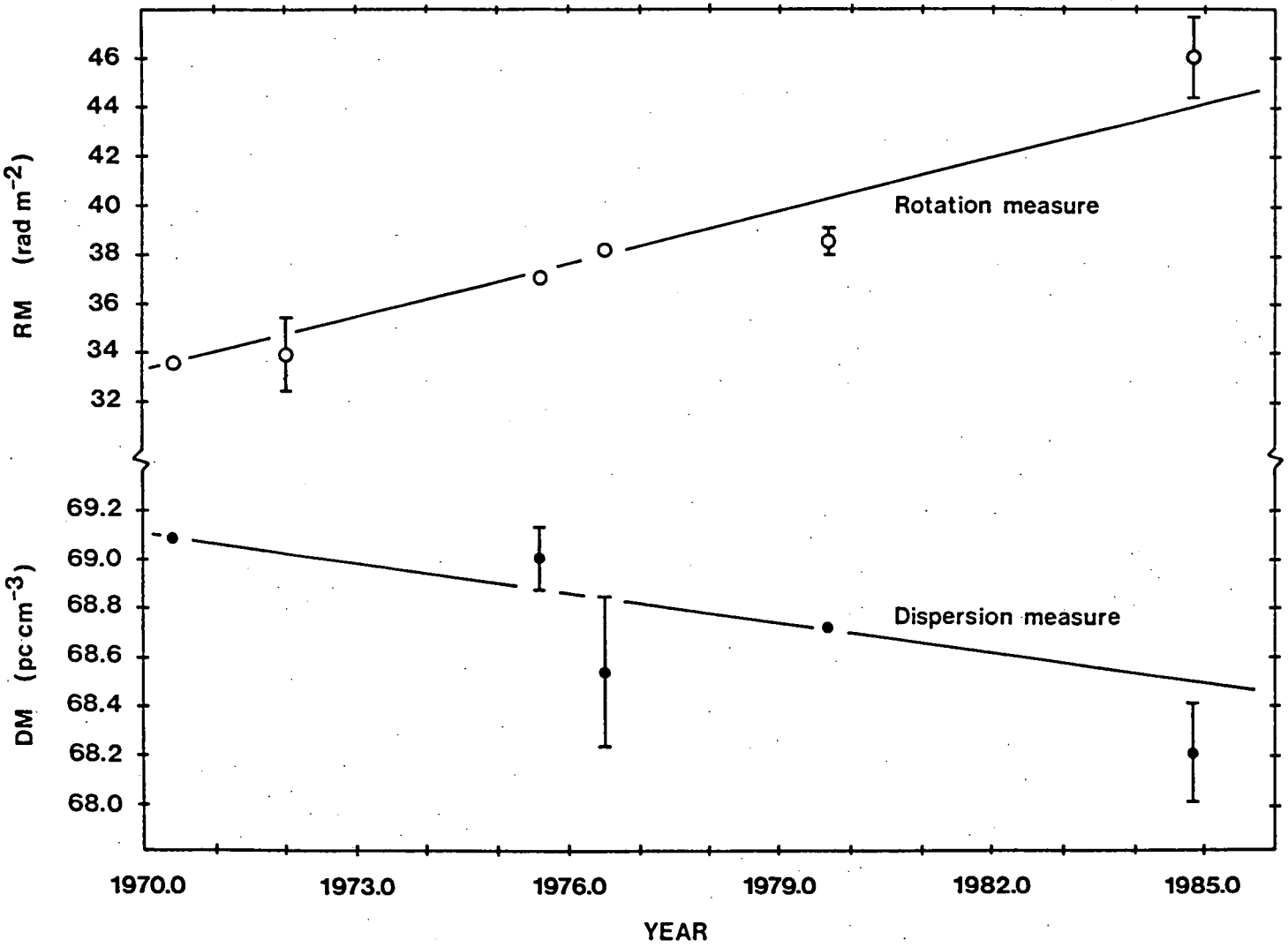


Fig.11-9: Dispersion measure and rotation measure of PSR 0833-45 plotted as a function of time. Error bars are ± 2σ. No error bars are shown if the quoted error would result in error bars smaller than the plotting symbol. The straight lines are derived using weighted linear regression. References are given in the text for data points recorded prior to October 1984.

where RM and DM are in the usual units and $\langle B_{\parallel} \rangle$ is in micro-gauss. With d(RM)/dt = 0.73 rad $m^{-2}yr^{-1}$ and d(DM)/dt = 0.04 pc $cm^{-3}yr^{-1}$, $\langle B_{\parallel} \rangle \sim 22$ μG. In normal interstellar regions B ~ 1 μG.

Supernova remnants often contain filaments of enhanced optical and non-thermal radio emission (see, for example, Cosmovici, 1974). The scale size of the postulated region of excess density together with the relatively high electron density and strong magnetic field suggests that the DM and RM changes may be caused by a magnetized filament moving across the line of sight to the pulsar. Values derived for the electron density and magnetic field strength are within the range found in supernova remnant filamentary structures.

The Hα study of the Vela remnant by Elliot et al. (1976) shows no intense filament near the pulsar but there are areas of diffuse nebulosity in the region. The diffuse nebulosity is brightest in the vicinity of PSR 0833-45 and the radio source Vela X. Elliot et al. suggest that the emission observed by them could be from a fossil Strömgren sphere produced when the local interstellar medium was ionized by radiation liberated during the supernova explosion. The results of the present study are consistent with the hypothesis that the nebulosity contains small-scale structure and regions where the magnetic field is enhanced. As well as the long-term DM and RM changes, there is a tendency for higher individual DM values (figure 11-9) to be associated with lower RM values and vice-versa. This may be an indication of even finer density and magnetic field structures within the obscuring region.

At low observing frequencies the pulse from the Vela pulsar shows a marked increase in duration. Ables et al (1970) and Komesaroff et al. (1972) have shown that this is due to scattering of the radiation by electron density inhomogeneities in the interstellar medium. There is a possibility therefore that the observed DM and RM changes are due to electron density and magnetic field variations in the scattering region.

Ables et al. (1970) show that the scattering screen is probably located at a distance of 250 pc, in the Gum nebula situated approximately mid-way between the pulsar and Earth. They estimate that the upper

limit on the scale size of irregularities in the region is 4×10^7 km (1.3×10^{-6} pc) and that the relative velocity of the medium is about 10 km s^{-1}, at least an order of magnitude less than the assumed pulsar velocity. The annual transverse displacement of the pulsar has been estimated (above) at about 5×10^{-4} pc, so DM or RM changes due to changing conditions at the scattering screen would be expected to occur on a time-scale of days rather than years. It is most likely therefore that the observed variations are the result of changing conditions in the Vela supernova remnant itself.

Continued monitoring of the long-term changes in the DM and RM of PSR 0833-45 is warranted from the point of view of extending the analysis undertaken in this section. As well, daily or weekly measurements may confirm short-term variations due to fine filamentary structure or scattering-related effects. These observations could be incorporated into the Llanherne timing experiment and would be best made using widely separated observing frequencies. Very accurate knowledge of the pulsar phase is a by-product of the timing experiment. Thus, separate polarization front-end units operating at different frequencies could be connected to the wideband polarimeter in synchronism with the pulsar frequency sweep. Alternatively, several continuously operating narrowband Stokes polarimeters could be employed. It would of course be of considerable interest to record polarization data during one of the occasional period "glitches" of PSR 0833-45. Linking the polarimetry and timing experiments would maximise the chance of this occuring.

CHAPTER 11 - REFERENCES

Ables, J.G., Komesaroff, M.M. and Hamilton P.A.: Pulse broadening of PSR 0833-45 by interstellar scattering. Astrophys. Lett., Vol.6, 1970, pp.147-150.

Cosmovici, C.B. (ed.): "Supernovae and supernova remnants". Reidel, 1974.

Elliot, K.J., Goudis, C. and Meaburn, J.: Deep Hα photography of the Vela and Puppis supernova remnants. MNRAS, Vol.175, 1976, pp.605-611.

Hamilton, P.A., McCulloch, P.M., Manchester, R.N., Ables, J.G. and Komesaroff, M.M.: Detection of changes in the rotation measure of the Vela pulsar. Nature, Vol.265, 1977, pp.224-225.

Komesaroff, M.M., Ables, J.G. and Hamilton, P.A.: Linear polarization measurements of PSR 0833-45 and their relevance to pulsar models. Astrophys. Lett., Vol.9, 1971, pp.101-104.

Komesaroff, M.M., Hamilton, P.A., and Ables, J.G.: Linear polarization and spectrum of PSR 0833-45 and the effects of scattering. Aust. J. Phys., Vol.25, 1972, pp.759-777.

McCulloch, P.M., Hamilton, P.A., Manchester, R.N. and Ables, J.G.: Polarization characteristics of southern pulsars - II. MNRAS, Vol.183, 1978, pp.645-676.

Rankin, J.M. and Counselman, C.C.: Pulsar NP0532: variability of dispersion and scattering. Astrophys. J., Vol.181, 1973, pp.875-889.

CHAPTER 12 CONCLUSION

This chapter contains suggestions for further signal processing research as well as a brief discussion of possible future modifications to the SAW coherent radiometer.

From the theoretical signal processing viewpoint there are several problems raised by the work presented in this thesis. Firstly, it would be satisfying to have analytical descriptions of the post-detection quantizing and integration processes. This may not be possible in the case of a linear detector where quantized Rayleigh variates must be summed. It is likely however that an analytical solution can be found for the square-law case. Secondly, an analytical model of the analog correlation and subsequent quantizing operation would be of interest. Studies by the author and Professor D.G. Lampard have so far failed to determine the video filter output p.d.f. as a function of correlation coefficient at the multiplier inputs. Quantization of the filter output is of course a separate problem. Finally, it would be useful to combine the processes of compression weighting, matched filtering, detection and video filtering in one model in an attempt to predict more exactly the effective bandwidth-time product in a SAW spectrum analyser resolution cell.

The most beneficial way of upgrading the prototype coherent radiometer would be to increase the duty cycle of the compressive receivers from 60% to 100%, giving a 1.1 dB increase in sensitivity. Bandpass filters at the SAW processor inputs, bi-linear signal mixers and a suitable pre-multiplier (LO) chirp generator would be necessary in this mode of operation (sections 4.4, 6.5, 6.6). The wideband analog multiplier described in section 6.7 is promising in the role of a bi-linear mixer for moderate bandwidth (< 55 MHz) analysers while a form of active chirp generation is one means of generating the required LO signal.

It is interesting to note that the existing passive LO generator will operate at 100% duty cycle. However, intermodulation effects in the non-linear LO chain together with overload in the signal mixer result in poor spectrum analyser performance.

The usable analysis bandwidth is restricted to about 15 MHz and the dynamic range is reduced significantly.

In military applications of compressive receivers, 100% processing duty cycle is sometimes achieved by using two separate analysers, each operating at 50% duty cycle and processing alternate spectral samples. However, the cost of SAW DDL's makes this "ping-pong" mode of operation expensive to implement. It is not the most attractive means of obtaining continuous operation of moderate bandwidth spectrum analysers.

In the prototype processor, the best way of obtaining 100% duty cycle operation would be to replace the existing SAW expansion devices with more suitable types. A single expansion DDL (TB~4 × 72) impulsed at intervals equal to half its chirp duration could be used as the source of an LO signal suitable for continuous operation. Either weighted or un-weighted convolvers (TB~72) could be used but the convolver must not be reciprocal ripple compensated. Replacement of the existing ripple compensated convolvers would improve greatly the processor sidelobe characteristics. Guidelines given in section 6.2 should be followed in choosing the new DDL's. The design example in section 6.3 is for a spectrum analyser having specifications similar to the prototype units but constructed around two DDL's of different time-bandwidth product.

In the longer term, replacement of the TDC1021J A/D converters with higher resolution or higher performance types would be worthwhile. Six-bit converters based on the same VLSI process as the TDC1021J devices have been tested but show the same deficiencies (section 8.9) as the four-bit converters. Samples of a new six-bit converter (TRW TDC1029) have been acquired and may prove satisfactory. The new devices have a maximum conversion frequency of 100 MHz and an analog section designed to handle full-scale 50 MHz signals without distortion.

The experience with commercial high speed A/D converters recorded in this thesis may be of interest to other researchers designing quantization systems for wideband noise signals. Manufacturers apparently assume Nyquist sampling in all applications.

Thus, when the analog section of a flash A/D converter is fabricated using the same VLSI process as the digital sections, the highest input frequency should not exceed half the maximum conversion frequency.

The digital and analog post-transformation processors developed do not relate specifically to the prototype SAW coherent spectrometer (TB = 72) but are adaptable to any SAW instrument with an output bandwidth of less than 30 MHz (or 55 MHz for the wideband correlator alone). The same basic designs are therefore applicable to SAW processors with TB >> 72. When wideband SAW spectrum analysers are used, it is possible to employ several 30 MHz DVI's operating in an interleaved processing mode. Alternatively, wideband analysis systems could be constructed by using radiometers having bandwidths of 30 MHz or so to process different parts of the spectrum. The second alternative may be attractive economically since the current component cost of the prototype coherent radiometer (SAW IF processor and four DVI's) is about $A8000.

The success of the initial scientific investigation augurs well for the future of the new equipment and techniques developed. In particular, the prospect of using the new instrumentation in conjunction with a large radio telescope is exciting. High time resolution polarization studies of galactic pulsars can begin immediately and one important application for the system in its high sensitivity mode is the measurement of the polarization behaviour of the extra-galactic pulsar, PSR 0529-66. For example, a 2.5 hour observation of PSR 0529-66 with the Parkes telescope should give a SNR of about 20, sufficient to allow the basic polarization properties to be deduced and an RM estimate to be obtained.

By necessity, this project has been a wide-ranging exercise in applied signal processing involving RF, high speed analog and fast digital processors. Many significant practical problems have been overcome and the SAW coherent radiometer is a functional scientific instrument. The highest priority must be given to the use of the equipment in this role.

APPENDIX I

Surface acoustic wave devices — fundamentals and applications

Surface acoustic wave (SAW) devices have been used for the last decade in professional and military equipment and are now beginning to be found in consumer electronic products. In these articles the basic physics of SAW devices are explained and some important current applications of the technology are examined.

P. J. Hall
Physics Department, University of Tasmania

Part 1

IN 1887 Lord Rayleigh, the great pioneer of modern acoustics, showed that a unique mode of wave propagation can exist at the surface of an elastic solid. This Rayleigh, or surface, mode wave has four important characteristics:

(1) The wave has a longitudinal component (solid displacement parallel to the direction of propagation) and a transverse component (displacement perpendicular to the propagation direction) each 90° out of phase. This gives rise to a backward elliptical motion of solid elements at the surface (Figure 1a). The amplitude of both wave components falls rapidly within the solid, confining the wave to within one or two wavelengths of the surface. The wave generates an oscillation of the surface as shown in Figure 1b.

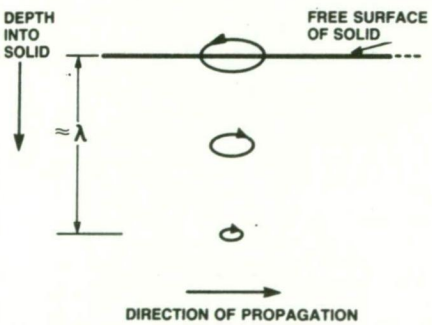
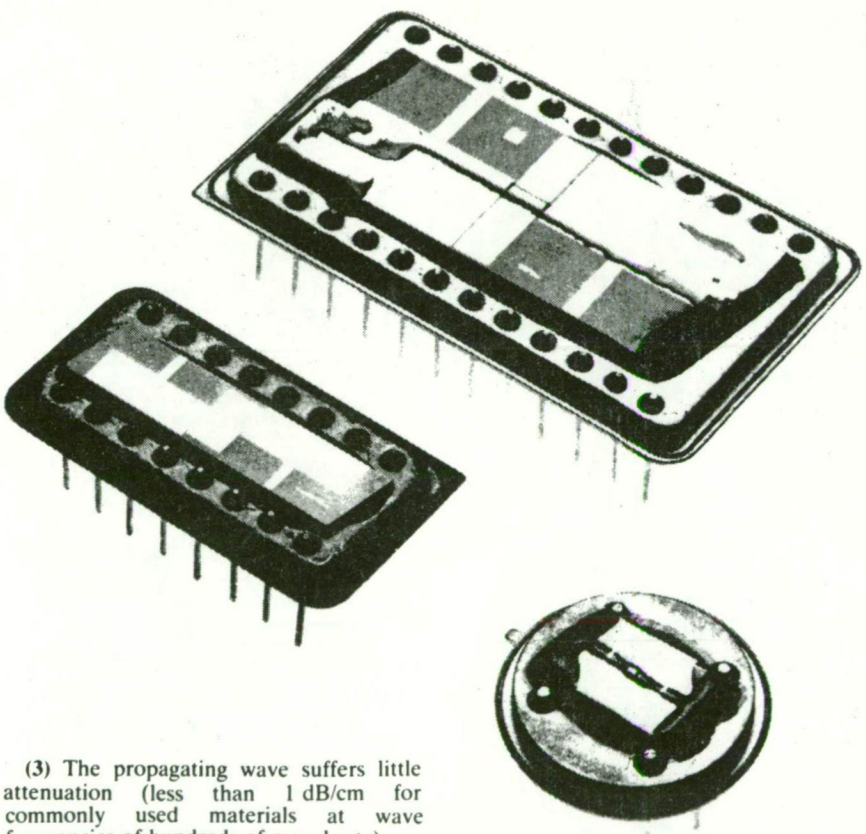


Figure 1(a). Motion of solid elements in an isotropic solid due to the propagation of a surface wave. Similar disturbances occur along axes of high symmetry in the anisotropic crystalline solids used for SAW device substrates.

(2) The wave velocity of propagation is of the order of 1500-4000 metres per second for most materials and is thus about 100 000 times as slow as electromagnetic wave velocities. The SAW wavelength is reduced by this factor, allowing a physically small length of solid to contain a great number of wavelengths.



(3) The propagating wave suffers little attenuation (less than 1 dB/cm for commonly used materials at wave frequencies of hundreds of megahertz).

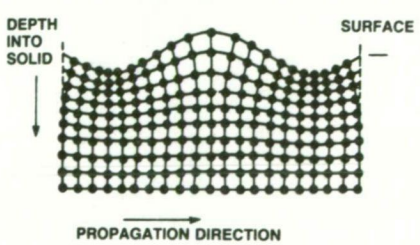


Figure 1(b). Oscillations in a vertical plane through a solid. Note the surface oscillations.

(4) The wave mode is essentially non-dispersive i.e: waves of different frequencies propagate with the same velocity.

The usefulness of SAW technology in electronics arises because it is possible to convert radio frequency electrical signals (over the range extending from a few megahertz to about one gigahertz) to surface acoustic waves and then back to electrical signals in a fairly efficient transduction process. ▶

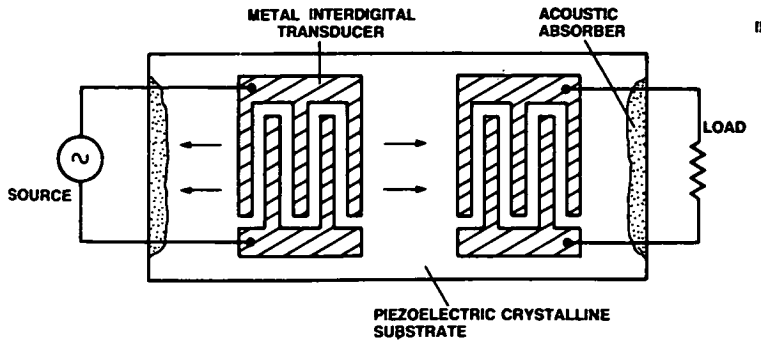


Figure 2. Basic SAW device configuration.

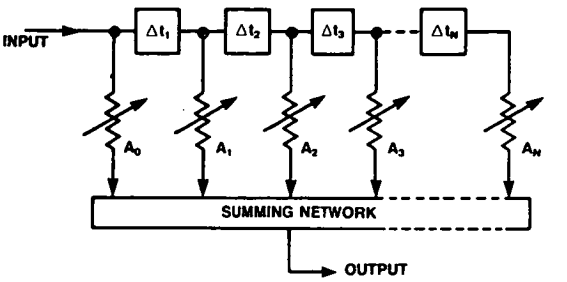


Figure 3. A transversal filter. In the SAW realisation the tap amplitudes A_0-A_N are set by the length of the IDT fingers and the time delay between taps is controlled by the relative positions of the fingers.

A mechanism for achieving this is shown in Figure 2. Using standard microelectronic fabrication techniques a metal interdigital transducer (IDT) is deposited on a piezo-electric crystalline substrate, usually quartz or lithium niobate. Although complex in detail the basic IDT operation is fairly easy to understand.

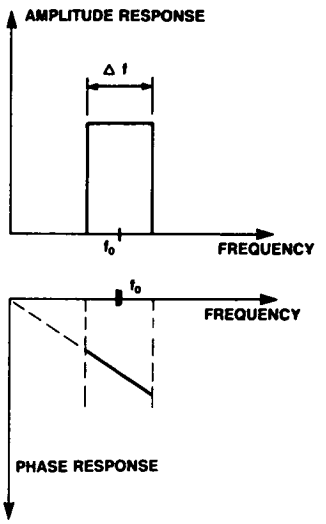
The electrical input signal is applied in such a way as to ensure that, at any instant, the charge on neighbouring IDT 'fingers' alternates in sign. The resultant electric field between the fingers leads to a mechanical stress in the piezo-electric solid and surface acoustic waves are launched in both directions. For maximum efficiency the physical spacing of the IDT fingers ought to be one half of an acoustic wavelength at the input frequency.

Wideband transducers can be designed by optimising the electrical and acoustic loads on the IDT. The propagating waves create electric fields in the solid which cause a voltage to be induced in the output transducer when the waves pass underneath.

The device illustrated in Figure 2 is an electrical delay line, the delay being simply the separation of the IDTs divided by the SAW velocity. Typical delays range from microseconds to hundreds of microseconds.

In many devices the unwanted energy propagated by the bi-directional IDT is either absorbed by an acoustic termination or scattered away from the functional surface. This leads to a minimum device insertion loss of 6 dB and also means that the portion of the signal reflected back from the receiving transducer, then back from the transmitter to finally arrive at the receiver IDT again, is suppressed by only 12 dB relative to the wanted output.

This 'triple transit' signal is the main device spurious response and can be further suppressed by deliberately mismatching the impedance of the IDTs, leading to an increased insertion loss. Fortunately the trade is reasonable: increasing the insertion loss from 6 dB to 12 dB improves the triple transit suppression from 12 dB to 33 dB.



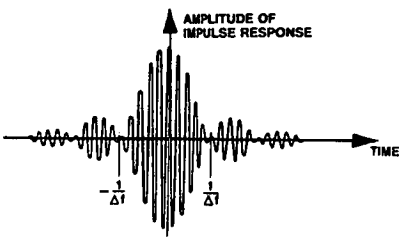


Figure 4. Responses of an ideal bandpass filter. The amplitude and phase responses are 'single-sided' (no negative frequencies) representations and the impulse response is derived by considering a psuedo-impulse (a very short pulse) modulated onto a carrier at the filter centre frequency. The impulse response is infinite in extent in time.

Polyphase IDTs are now available which launch a unidirectional surface wave and lead to devices with much lower insertion loss and better triple transit suppression. Nevertheless, the relatively high insertion loss of most currently available SAW devices is due to IDT mismatch.

To appreciate the signal processing potential of SAW devices it is necessary to look on the IDT as a sampled delay line. Surface waves travel from one pair of electrodes to the next at a finite velocity so the input signal is repetitively delayed and added to itself. The constructive or destructive signal addition creates the amplitude and phase response of the SAW device.

The amplitude of a particular delay line tap can be set by choosing the length of the IDT fingers at that point. The phase, or time delay, of the tap can be varied by changing the relative position of the fingers. In a given SAW device either the transmitter or receiver IDT, or a combination of both, may be used as the signal processing element.

It turns out that such a delay line is exactly what is required to form the basis of a transversal filter. In such a device a given frequency response is obtained by synthesising the time domain impulse response of the desired filter. The impulse response is the output from the filter when the input is a very short pulse and is, in fact, the Fourier transform of the desired frequency domain characteristic.

Figure 4 shows the impulse and frequency response of an ideal band-pass device, the type of filter to be discussed later.

The impulse response synthesized by a transversal filter is necessarily finite in length and sampled rather than continuous, but by following the usual principles applying to the design of sampled systems, the desired impulse response can be well approximated.

To use a SAW device as a filter it is only necessary to 'draw' (using metallization) the sampled impulse response on the substrate. The second IDT can be a broad-band type designed to pass all frequency components of interest (Figure 5).

One major constraint applies to all SAW filters. Since the polarity of the charge must reverse on alternate IDT electrodes in order to establish surface wave propagation, the sign of alternate samples of the impulse response reverses.

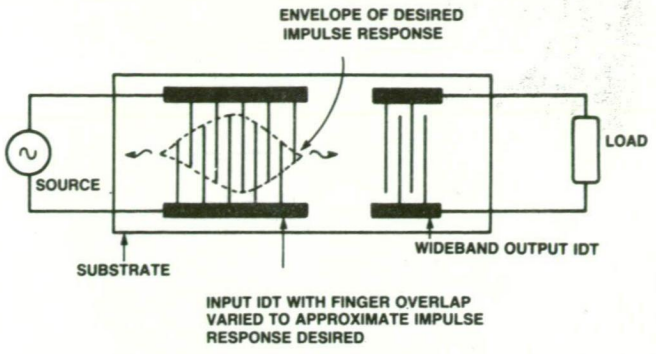


Figure 5. SAW transversal filter.

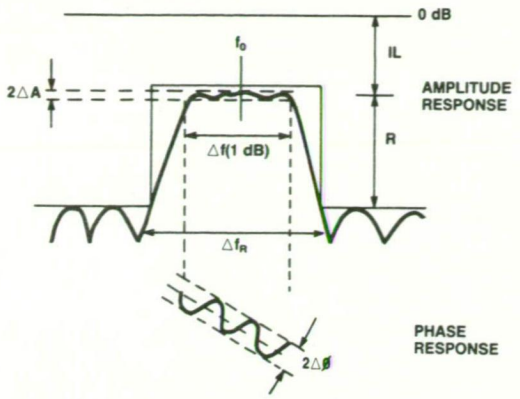


Figure 6. Major parameters of a bandpass filter.

Sampling in this way produces a pass band displaced from baseband i.e: a bandpass filter.

A major advantage of SAW filters is that they are not 'minimum phase' devices, therefore the amplitude and phase responses can be specified independently. Hence, a fairly sharp frequency response can be combined with a linear phase response, often important in video and data distribution networks.

Before examining some specific applications of SAW filters it is worth summarizing some of the advantages of the technology. These advantages can be listed as:

(1) Versatile response with essentially no adjustment or tuning.
(2) Straightforward fabrication using standard techniques.
(3) Excellent repeatability from device to device due to the photo-lithographic processes used in making the IDTs.
(4) Small size and weight. Mechanically rugged.
(5) Passive.
(6) Wide frequency range of operation (10 MHz to 1 GHz).
(7) Good temperature stability.
(8) Graceful degradation (small faults cause small performance changes).
(9) Radiation resistant.

Table 1. Typical Parameters of SAW transversal bandpass filters

Substrate material	Lithium niobate or quartz
Centre frequency (f_0)	10-1000 MHz
1 dB bandwidth (Δf)	60 kHz to 50% fractional bandwidth
Insertion loss (IL)	10-30 dB
Amplitude ripple (ΔA)	± 0.2 dB
Phase ripple ($\Delta\phi$)	± 1°
Ultimate rejection (R)	50-70 dB
Shape factor ($\Delta f_R/\Delta f$)	1.1
Size	10-30 mm

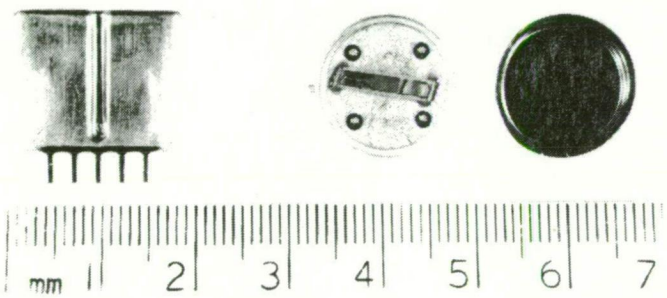


Figure 7(a). Two versions of a SAW TV IF filter. The filter on the right has the top removed to reveal the substrate and IDTs (devices courtesy of Signal Technology Ltd).

SAW bandpass filters

The primary specifications for a bandpass filter are the centre frequency, f_0, and the bandwidth, Δf. These and other secondary specifications are explained in Figure 6. It is convenient to express the bandwidth in a normalized form, the percentage franctional bandwidth of the filter being simply $(\Delta f/f_0)x100$.

Table 1 details the achievable performance with typical modern SAW filters but does not include data for the new devices with polyphase IDTs.

The most commonly used substrate material is lithium niobate, but in filters with a small fractional bandwidth (less than 5%) quartz may be used to ensure greater temperature stability. The centre frequency range is restricted at the low frequency end by large device die sizes and a poor spurious response performance by the filter. At the high frequency limit, photo-lithographic processes used to manufacture the IDTs are severely taxed.

The filter bandwidth limitations are set essentially by the number of fingers in the IDT. For very narrow bandwidths (less than 60 kHz) the number of fingers needed is very large, causing an intolerable number of reflections and spurious responses.

In situations where such narrow bandwidths are needed, a second type of SAW filter is used. These 'resonator' filters are the SAW analogy of microwave cavities and can achieve fractional bandwidths of 0.01% or so.

At the other extreme, fractional bandwidths of more than 40% are difficult to synthesize with SAW transversal filters because of the very small number of IDT fingers employed to approximate the wanted response.

The insertion loss figures quoted in Table 1 are typical but the adoption of polyphase IDT techniques can reduce the insertion loss to as low as 1.5 dB. As a general rule, increasing the fractional bandwidth of a SAW filter increases the insertion loss since the resistive loading of the IDT must be increased.

Amplitude and phase ripple in a well-designed filter are due mainly to spurious device responses. These responses also govern the ultimate rejection capability of the filter. The 50-70 dB rejection quoted is achieved routinely in modern filters.

Figure 7a shows a typical SAW bandpass filter, in this case a 38.9 MHz TV IF filter. The same device is available in a choice of the two packages shown and both encapsulations sell for a few dollars.▶

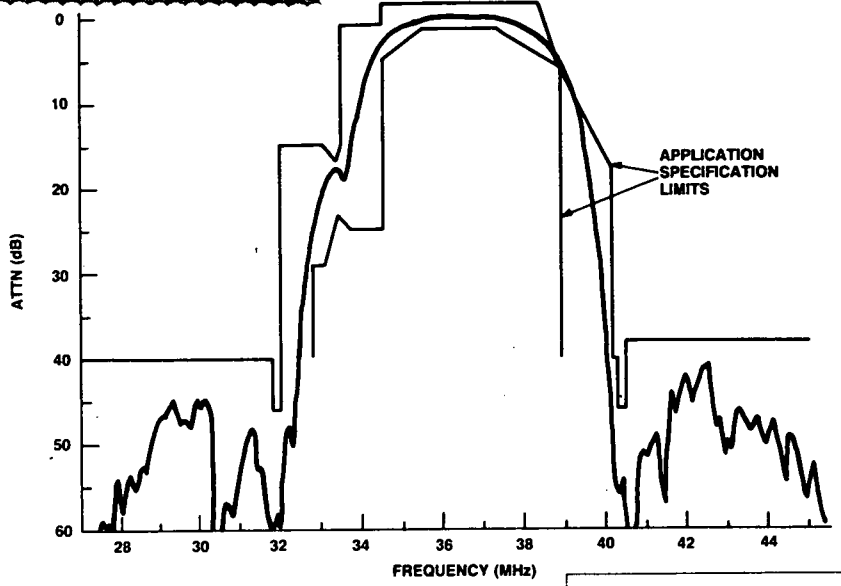


Figure 7(b). Response of TV IF filter (courtesy of Signal Technology Ltd).

The frequency response of the filter is shown in Figure 7b and is close to the ideal response for the TV application.

Most SAW devices are not produced in the same numbers as TV filters but nevertheless the price is attractive in many situations, especially when the total cost of alternative filters is considered.

In addition to the advantages of high performance, no alignment filtering, SAW devices offer exceptional device reproducibility. Figure 8 demonstrates the exceedingly small variation in response expected in production SAW filters. This advantage is present in all SAW devices and is often of great benefit in coherent (phase sensitive) communication and signal processing systems.

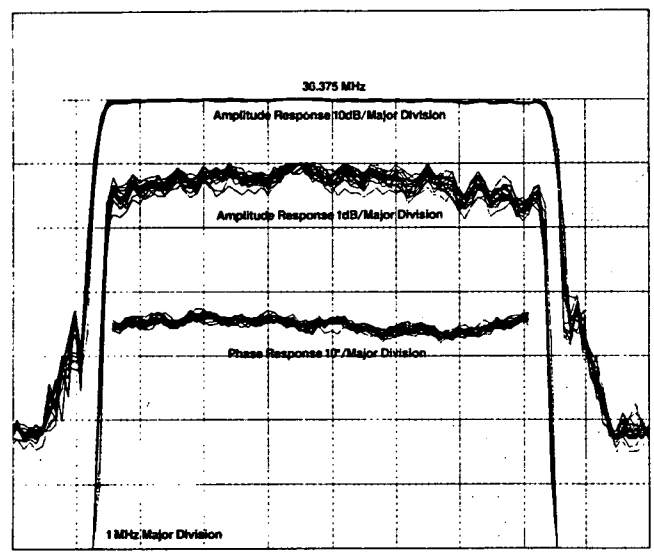


Figure 8. Overlay of the responses of 20 SAW filters (courtesy of Anderson Laboratories Inc).

Time coded filters

When examining SAW bandpass filters it is convenient to retain a conventional frequency domain specification of the filters even though the response is actually synthesised in the time domain. It is also stressed that SAW filters may have independent specifications for the phase and amplitude responses.

Bandpass filters are normally designed with a linear phase characteristic or, equivalently, with a constant group delay (i.e: all inband frequencies entering the filter emerge at the same time). A device that exhibits a flat group delay is said to be 'non-dispersive'. In fact, there are applications where dispersion (the emergence of different frequencies at different times) is required and where SAW dispersive filters are now used extensively. These devices are best characterized by their time domain rather than their frequency domain responses.

Chirp radar

One of the most common professional uses of SAW technology is in pulse compression or chirp radar systems. An impulse, which in theory contains an infinite number of frequency components, is applied to a SAW device known as a dispersive delay line or DDL (Figure 9). The DDL output is a dispersed version of the input and is a frequency modulated pulse or chirp centred at f_0 and extending over a range B in frequency and T in time.

Note that the dispersive nature of the device arises because of the non-uniform spacing of the IDT fingers rather than as a consequence of any dispersion during propagation. The chirp slope of the DDL is simply B/T and the device time-bandwidth product, N=BT, is an important parameter.

In the radar application (Figure 10) the DDL output is amplified, transmitted, reflected from the target and the returned echo applied to the receiver. The heart of the receiver is a second DDL, the dispersion characteristic of which is the exact inverse (i.e: chirp slope of −B/T) of the transmitter DDL.

The effect is to compress the returned signal back to a pulse resembling the transmitter impulse, except for 'side lobes' due to the finite bandwidth of the expansion-compression process. The time between the excitation impulse and the appearance of the compressed pulse is directly related to the range of the object in the radar beam.

This may appear to be an unnecessarily complex scheme but in practice a major advantage emerges. In all radar systems the detection sensitivity is proportional to the transmitted energy and the time resolution is proportional to the transmitted bandwidth.

In real radars the peak transmitter output cannot be increased indefinitely and once the limit is reached the only way to increase the energy output of a normal pulsed radar is to lengthen the pulse. Unfortunately, the bandwidth of the pulse is then decreased and the time, and therefore the range, resolution is reduced accordingly.

In a chirp radar the transmitted pulse bandwidth is artificially increased by the superimposed frequency modulation. A pulse long enough to ensure adequate detection sensitivity can be used while at the same time retaining sufficient bandwidth for good time resolution.

It is found that the time resolution of a chirp radar is the actual pulse length, T, divided by the time-bandwidth product, N=BT. Since N can be large (>1000 in many cases), the advantage is real. Note

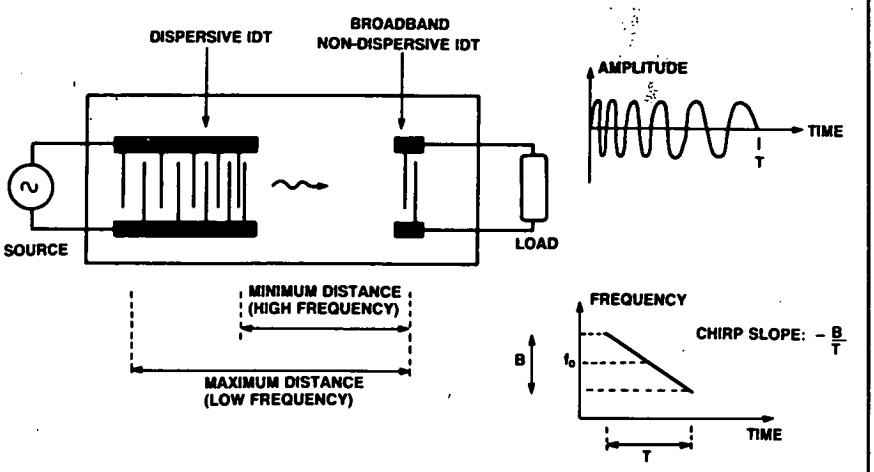


Figure 9. A dispersive delay line with two representations of its impulse response. When an impulse excites the dispersive IDT, the high frequency components arrive at the output before the low frequency components. The resultant output is a 'chirp' sweeping down in frequency. The complementary 'un-chirp' can be generated by reversing the high and low frequency ends of the dispersive IDT.

that since T/BT = I/B, the compressed pulsewidth or time resolution is of the order of the reciprocal bandwidth.

The side lobes of the compressed pulse may be a problem in some radar applications since a weak echo can be obscured by the side-lobes of a nearby strong echo. It can be shown that if the frequency spectrum of the chirps is uniform, the first side-lobe is down only 13 dB with respect to the peak of the compressed pulse.

A partial solution is to taper the chirp spectrum of the receiver DDL impulse response, a process known as 'weighting' the chirp. The result is a broader compressed pulse exhibiting side-lobes down 40 dB or so.

SAW technology has been embraced enthusiastically by radar designers and most major SAW device manufacturers sell DDLs (weighted and unweighted) primarily intended for radar use. The ruggedness of the devices makes them ideal for use in environments such as aircraft, ships and other surroundings usually considered hostile to electronic components.

Before returning to SAW DDL's in their second major application, spectrum analysis, it is worth expanding on the concept of matched filtering, of which pulse compression is a particular example.

SAW devices as matched filters

If a signal is written as a function of time, denoted by s(t), then the impuse response of a filter 'matched' to the signal is s(−t) i.e: a time-reversed form of the signal. The output from such a filter is known as the auto-correlation function of the input signal and is the output exhibiting the best signal-to-noise ratio obtainable in the presence of white noise.

In general, a matched filter cannot be built because causal principles would need to be violated. The output would need to precede the input but, in practice, excellent approximations are possible.

Referring to the chirp radar system in Figure 9, it is clear that the receiver DDL is a matched filter for the transmitter chirp since the compressor impulse response is the same as the expander chirp reversed in time. The classic (sin x)/x compressed pulse can indeed be shown to be the auto-correlation function of the transmitted chirp.

SAW matched filters can of course be designed for any waveform and such filters will undoubtedly find application in both radio and line circuits used for the transmission of digital data. Already SAW devices are being used as code generators and correlators in modern spread spectrum communication systems. ●

. . . to be continued.

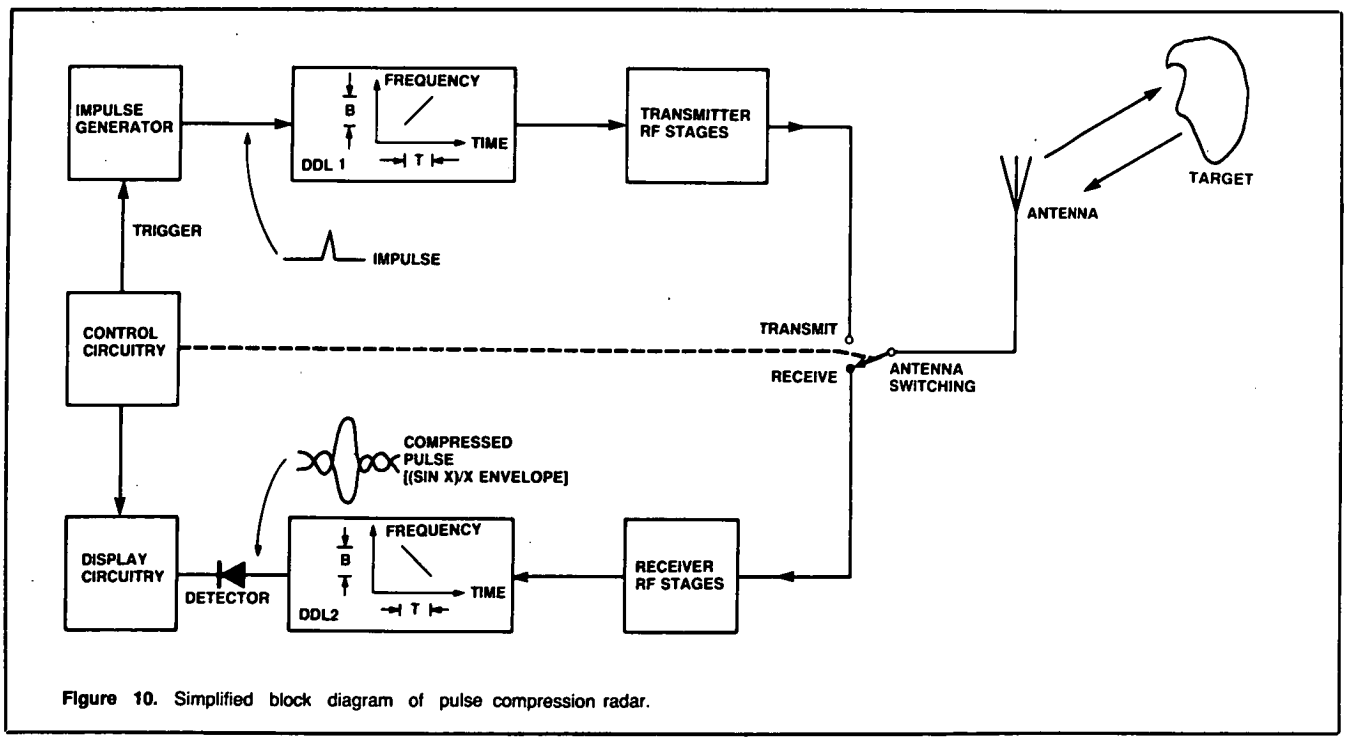


Figure 10. Simplified block diagram of pulse compression radar.

Surface acoustic wave devices — fundamentals and applications

The basic physics of SAW devices was examined in part 1. In this article a major application of time coded SAW filters, spectrum analysis, is explained.

P. J. Hall
Physics Department, University of Tasmania

Part 2

THE THIRD MAJOR use of time-coded SAW filters is in spectrum analysis and, as in chirp radars, the basic component is the dispersive delay line (Figure 10).

Before examining SAW spectrum analysis methods, it is useful to consider alternative techniques and their limitations. This section is mainly concerned with analysis systems designed for on-line signal processing applications. In such applications, spectrum analysis is used to obtain information about signals on a real-time 'as the action happens' basis, perhaps in order to control some other process e.g: the electrical steering of an antenna array to minimise interference to a radio link.

The most obvious way of analysing a given bandwidth is to assemble a bank of frequency contiguous filters covering the analysis bandwidth and to relate the output of each filter to the spectral energy falling within the passband of that filter. Filterbanks are commonly used in audio, radio astronomy and some military applications with refinements dependent upon the specific use.

Unfortunately, when a large bandwidth is to be analysed with good frequency resolution, the number of filters becomes prohibitively great and the analyser becomes large, expensive, difficult to align, and difficult to keep in alignment. Filterbanks consisting of SAW bandpass filters are now available and offer an excellent alternative to LC filters in some RF applications, although in most situations the more elegant SAW techniques to be described would normally be used.

Spectrum analysis

Conventional scanning spectrum analysers are acceptable for laboratory use but are too slow and too wasteful of signal energy to be used in most signal processing roles. A signal shorter than the time the analyser takes to scan the analysis bandwidth can come and go without ever registering on the instrument.

In military jargon, the analyser has a very low probability of intercept (POI). If the scan rate is increased to raise the POI, the available frequency resolution is decreased. The scanning analyser wastes energy because if 'N' frequency resolution cells are resolved across the analysis bandwidth, the time spent actually looking at each cell is only 1/Nth of the total observing time.

Digital spectrum analysers based on the Fast Fourier transform (FFT) algorithm are

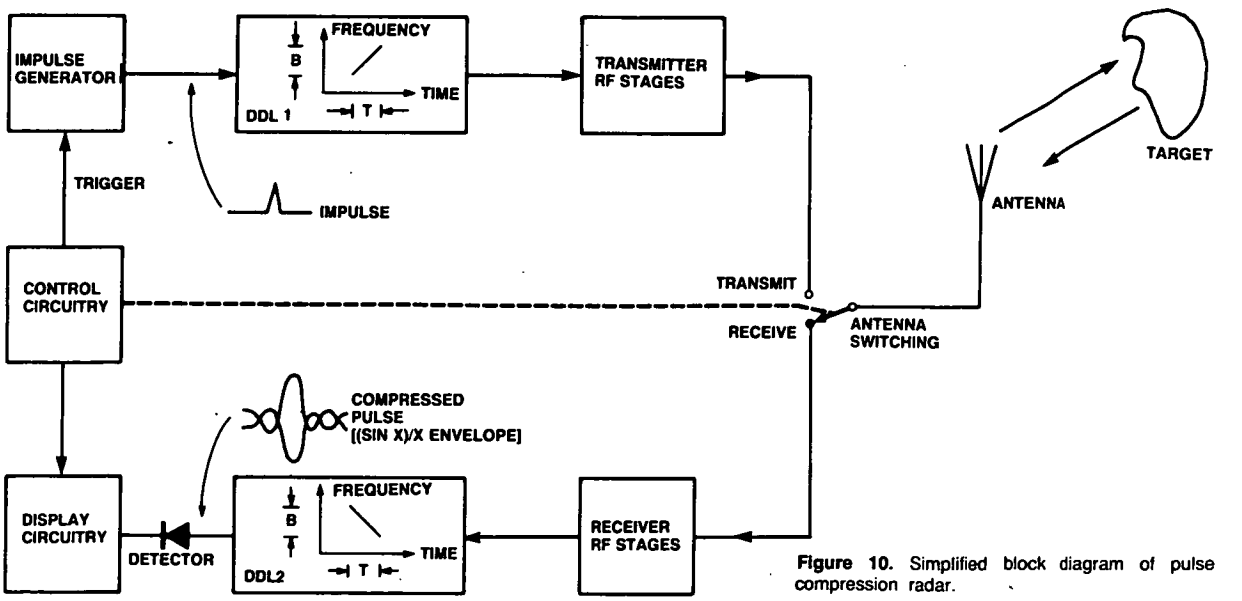


Figure 10. Simplified block diagram of pulse compression radar.

now familiar in audio test and development laboratories as well as in some signal processing applications. FFT methods produce accurate spectra and are capable of conveniently providing a true Fourier transform i.e: phase as well as amplitude, or power of spectral components. Unfortunately, the analysis bandwidth for real-time FFT processors is currently rather limited with most commercial units being essentially audio analysers.

A second digital method of power spectrum analysis, the auto-correlation technique, can be extended to wide bandwidths but is not suitable as a general approach, particularly in applications where high time resolution i.e: large number of spectra per second is needed.

Another instrument sometimes used for spectrum analysis in radio astronomy and military applications is known as the acousto-optical spectrum analyser. This device relies on the interaction of laser light with RF sound waves in a translucent piezoelectric crystal. Wide bandwidths can be analysed with good frequency and time resolution but current designs are plagued with a number of practical problems, the most serious being mechanical and thermal instability and limited dynamic range.

Coherent (phase sensitive) acousto-optical processors have been described but with existing technology, practical problems in such systems are acute. Integrated acousto-optical processors which utilise the interaction of light from a solid state laser with surface acoustic waves may substantially avoid the problems of existing optical analysers.

SAW dispersive delay line spectrum analysers overcome many of the disadvantages of the other types of instruments and offer a state-of-the-art solution to the problem of real-time, wideband Fourier analysis.

Compressive receiver

A block diagram of a SAW 'compressive receiver' spectrum analyser is shown in Figure 11. This configuration, which is only one of several possible arrangements, uses two SAW dispersive delay lines (DDLs) similar to the type used in the chirp radar system described earlier.

One DDL is used to expand an impulse into a chirp, and this chirp is then mixed with signals in the band to be analysed. Consider CW input signals for simplicity. Each signal produces a chirp from the mixer which is time co-incident with, but offset in frequency from the other chirps. This is illustrated in the frequency-time diagram for the process (Figure 12).

Signals spaced in frequency at the input produce chirps which enter the passband of the second DDL, the convolver, at different times. The chirp slope of the convolver is the exact inverse of the expander DDL chirp slope so chirps entering the convolver emerge as compressed pulses, just as in the pulse compression radar. In this case however, chirps due to different input frequencies enter the convolver at different times, so the compressed output pulses are separated in time.

Hence, separation in frequency at the compressive receiver input produces separation in time at the output, so the instrument functions as a spectrum analyser. The time axis on a display device e.g: an oscilloscope can be read as frequency by using the convolver chirp slope as a scaling factor, i.e: $f = (B/T)t$. This process is an example of the 'chirp transform' and, as mentioned earlier, other practical implementations are possible.

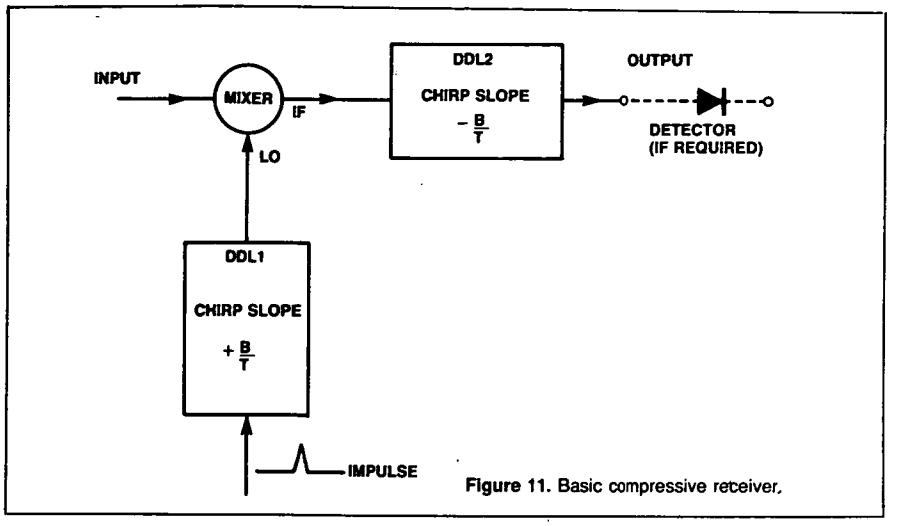


Figure 11. Basic compressive receiver.

Several key characteristics of the compressive receiver are related to the time-bandwidth product, $N = BT$, of the convolver. As expected, the width in time of the compressed pulses is of the order of the reciprocal bandwidth, 1/B. The receiver takes T seconds to analyse the total bandwidth, B. The number of identifiable frequency resolution cells is therefore about $T/(1/B)$ or N, so the time-bandwidth product is analogous to the number of channels in a filterbank or the number of points in a digital FFT analyser. The equivalent width in frequency is simply B/N or I/T.

It can be shown that if a compressive receiver is designed to give a frequency resolution equivalent to a scanning spectrum analyser, the compressive receiver can analyse the input bandwidth a factor of N faster than the scanning analyser. Since N can be up to 10 000 with currently available SAW devices, the advantage can be very large. As shown in Figure 12, it is usual to use an expander chirp with twice the bandwidth and duration of the convolver impulse response.

This ensures that chirps due to signals at either edge of the analysis bandwidth emerge from the convolver fully compressed, but it also means that chirps due to some very short duration input signals can fall outside the convolver bandwidth for part of the active cycle of the receiver. ▶

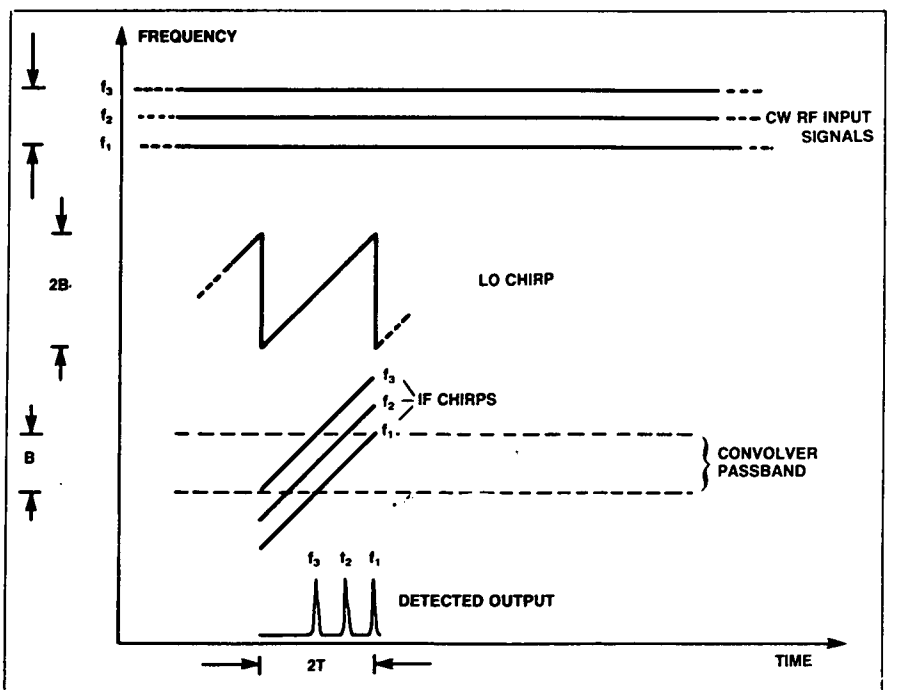


Figure 12. Frequency-Time diagram for compressive receiver. Note the duality between frequency at the input and time at the output.

The effect is to lower the POI to 50% for signals of duration less than T (typically microseconds). Nevertheless, this figure is much higher than the POI factor for almost all other analysers. A high POI for SAW compressive receivers makes the instruments popular with military users, who often use more advanced techniques to raise the POI to 100%.

SAW analysers

When compared with digital spectrum analysis techniques, the most obvious advantages of SAW analysers are high speed and wide analysis bandwidth. It has been estimated that SAW analysers exhibit a speed advantage of from 100 to 10 000 depending on the DDLs used and the type of digital instrument used as the basis for comparison. In certain applications the overall benefit of adopting SAW technology can be very great.

For example, the authors of a U.S. paper describing a sophisticated SAW signal processing module claim an advantage factor for SAW technology over digital methods of 12 000. This figure is based on a system figure of merit defined as 'the equivalent number of complex multiples per second per dollar of development cost'.

SAW processors are similar to digital instruments in that both are capable of providing amplitude and phase spectra and, as such, are useful in roles requiring coherent signal processing. Part of the recent work in radio astronomy at the University of Tasmania has been to demonstrate that coherent SAW spectrum analysers are viable now, using existing technology and techniques.

In contrast, coherent optical methods (the main rival to SAW technology for wideband, real-time applications) are some way from being fully developed. As a bonus, SAW technology offers a dramatic improvement over optical methods in the areas of mechanical and thermal stability, and dynamic range.

The limitations associated with SAW spectrum analysers are fairly minor. The analog transformation process exhibits an accuracy equivalent to a 6 or 7 bit digital transform. For many applications this is sufficient, particularly when the speed advantage of a SAW instrument is considered.

The second practical problem is that SAW analysers are simply too fast to be directly useful in applications where the instruments need to be interfaced to computers or other digital post-transformation devices. For example, two University of Tasmania spectrum analysers output data at a rate of around 40 ns per frequency resolution cell, and this is the rate at which a digital processor needs to sample the detected output.

We have solved the problem by producing 'digital video integrators', or DVIs, to act as an interface between the SAW analysers and slower logic. A DVI samples the analyser output at high speed and adds together many spectra before the integrated spectrum is output to a computer or other device. The DVI runs continously so there is no loss of new data while the integrated spectrum is output.

This process degrades the time resolution slightly but, even so, the integrated spectrum time resolution can approach 100 μs with existing data acquisition computers. Of course, the wide-band, coherent processing capability of SAW spectrum analysers is retained.

The DVI techniques developed in Tasmania have created interest overseas, and it is hoped that with the local VLSI program underway, further work can be undertaken on the design of an Australian produced parallel processor ship for use in the back-end of a DVI.

Conclusion

These articles have summarised some important current uses of SAW technology but the treatment has been necessarily brief. Discussion of important devices such as reflective array compressor (RAC) dispersive delay lines, non-linear acoustic convolvers, SAW oscillators and SAW memories has been omitted.

Interested readers are referred to the general references listed below. For the benefit of designers considering using SAW devices I have provided a list of suppliers known to me. Most companies can supply customised as well as standard stock devices.

GENERAL REFERENCES

1. Butler, M.B.N., 'Surface acoustic wave devices', *Electronic Engineering*, June 1980, July 1980, February 1981.
2. Dieulesaint, E. and Royer, D., *'Elastic waves in solids'*, Wiley, 1980.
3. Hartmann, C.IS. et al., 'SAW devices for military, communications, radar and EW systems', *Microwave Journal*, July 1982.
4. Oliner, A.A. (Ed.), *'Acoustic surface waves'*, Springer-Verlag, 1978.
5. Penimuri, D. and Havens, D.P., 'Surface acoustic wave filters prove versatile in VHF applications', *Electronics*, July 5, 1979.

SAW DEVICE SUPPLIERS

1. Andersen Laboratories Inc., 1280 Blue Hills Avenue, CT 06002, USA. TWX 710-425-2390.
2. Crystal Technology Inc., 1035M E. Meadow Cir., Palo Alto, CA 94303, USA.
3. GEC — Marconi Research Centre, Great Baddow, Essex, CM2 8HN, U.D. Telex 99201.
4. Hughes Aircraft Company, Gound Systems Group, P.O. Box 3310, Fullerton, CA 92634, USA.
5. Racal-MESL Ltd., Lochend Industrial Estate, Newbridge, Midlothian, EH28 8LP, U.K. Telex 72384.
6. RF Monolithics Inc., 4441 Sigma Road, Dallas, Texas 75234, USA.
7. Siemens Ltd., Australian Office, 544 Church Street, Richmond, Victoria, 3121. Telex AA30425.
8. Signal Technology Ltd., Cheney Manor, Swindon, Wiltshire, SN2 2PJ, U.K. Telex 444410.
9. Thomson-CSF, Chemin des Travails, B.P. 53, 06802 Cagnes-Sur-Mer, France. Telex 204.780F. ●

APPENDIX II - SAW DISPERSIVE DELAY LINES

The basic design concepts and main applications of surface acoustic wave (or Rayleigh wave) devices have been discussed in appendix one, a reprint of a review article by the author. This appendix deals specifically with the design and performance of SAW dispersive filters.

A basic SAW DDL is shown in figure I-9 (appendix one) and in figure II-1. The filter is essentially a delay line since Rayleigh waves are launched from the sending interdigital transducer (IDT) into the non-dispersive piezo-electric propagation medium and then converted back to electrical signals by a receiver IDT. The output from the receiving transducer appears delayed by a time $t_d=d/v_o$ relative to the input signal, where d is the acoustic path length and v_o is the Rayleigh wave velocity in the crystalline substrate (about 3000 ms^{-1} for commonly used materials).

The IDT launches Rayleigh waves most efficiently at input frequencies near the synchronous frequency: the frequency where the transducer finger spacing is one half of an acoustic wavelength. By varying the synchronous frequency along the transducer length it is possible to make a dispersive filter. In such a device electrical signals containing many frequency components have these components converted selectively into acoustic waves at different parts of the IDT. Since the propagation path length differs for each frequency, the time of arrival of each component at the output transducer is different. In other words, the output is a time dispersed version of the input. In practical devices the input or output IDT (or both) may perform the dispersion.

If a dispersive filter is excited with an impulse the output is a dispersed impulse or FM pulse, the time and frequency truncation characteristics of which are determined by the IDT. The impulse response of the filter is therefore qualitatively of the form specified by equation 3.2.1, so the SAW DDL can be used as an expander to generate chirp signals or as a compression filter.

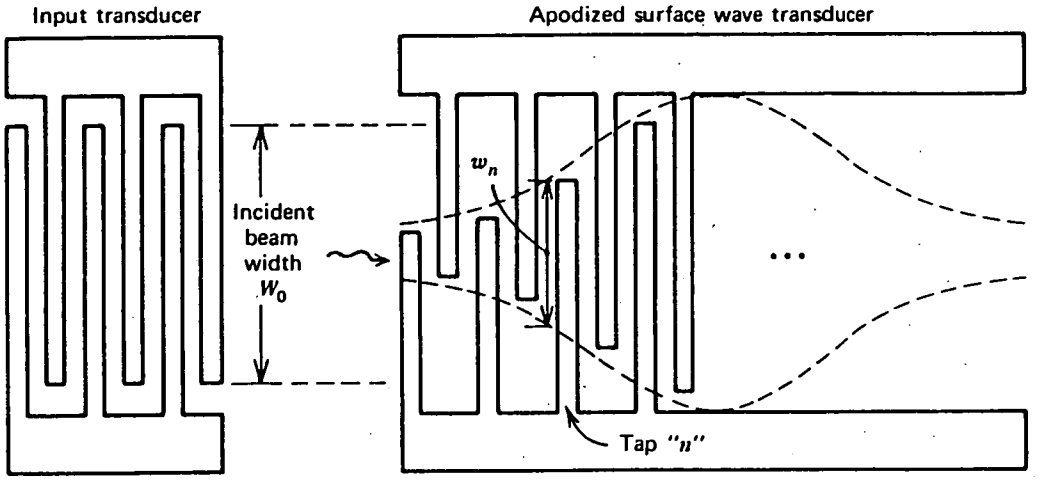


Fig.II-1: SAW chirp filter (After Gerard, 1977).

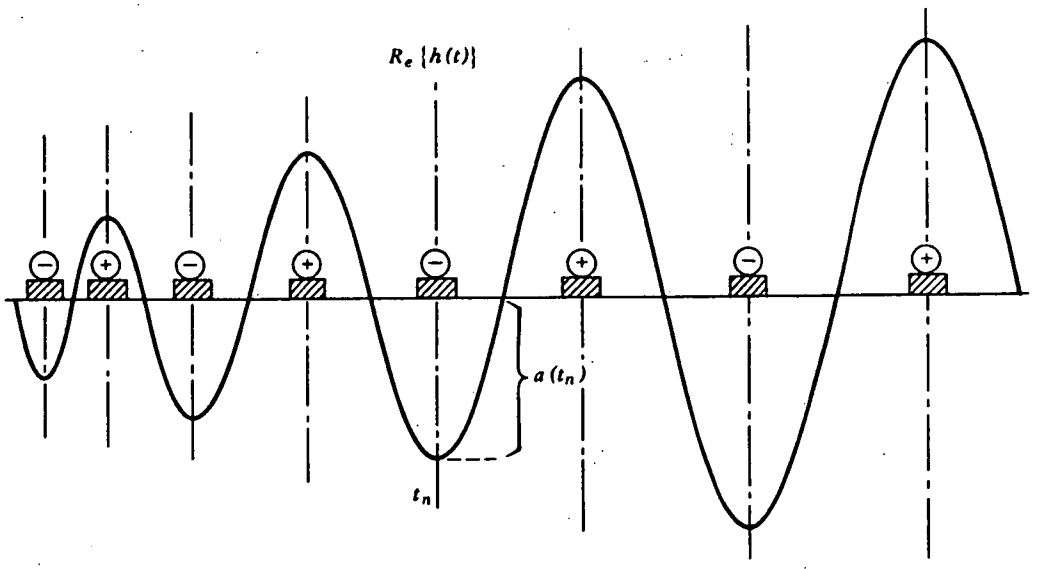


Fig.II-2: Electrode positions for sampling a signal at times for which the sample weights are real. (After Gerard, 1977).

To synthesize a SAW filter capable of generating an arbitrary impulse response (which is the Fourier transform of a given frequency response) it is necessary to view the IDT as a transversal network, the gaps between the fingers corresponding to the taps and the substrate providing the inter-tap delay. A given impulse response is synthesized by providing taps with amplitude and phase weights corresponding to the values obtained when the impulse response is sampled in accordance with the requirements of normal sampling theory. Amplitude weights can be introduced by means of apodization (variation in the length of the IDT fingers) and phase weights are set by the position of the fingers on the substrate.

Figure II-1 shows that, because of the bus summing arrangement, alternate samples must have a polarity of ± 1 or a phase weight of $n\pi$ radians (Gerard, 1978). In practice, the fingers are positioned at points corresponding to sampling times at which the required impulse response is real (figure II-2).

To produce a general impulse response

$$h(t) = w(t)\exp[j\phi(t)]$$

with a frequency-time characteristic given by

$$f(t) = \frac{1}{2\pi}\frac{d\phi}{dt} \ ,$$

the sample times, t_n, for the transversal network are found by solving

$$\phi(t_n) = n\pi + \text{constant}, \qquad \text{II.1}$$

where n is an integer (Gerard, 1977). The corresponding spatial positions of the IDT electrodes are given by

$$z_n \sim v_o t_n \ , \qquad \text{II.2}$$

where v_o is the free-surface wave velocity. The effective velocity (v_{eff}) differs from v_o because of the effects of piezo-electric "shorting" by the metal IDT electrodes and of acoustic mass loading by the metallization.

If the IDT pattern is such that electrodes and adjacent gaps have equal widths then

$$z_n = v_{eff}\, t_n$$

and

$$v_{eff} = v_o/[1 + \tfrac{1}{2}\tfrac{\delta v}{v_o}] \sim v_o(1 - \tfrac{1}{2}\tfrac{\delta v}{v_o}) \quad , \qquad \text{II.3}$$

where $\delta v/v_o$ is the fractional change in velocity caused by the presence of the metallization. A reasonable approximation is $\delta v/v_o \sim K^2/2$, where K^2 is the piezo-electric (or electro-mechanical) coupling constant of the substrate (Farnell, 1977).

Consider now the electrode pattern required to produce the uniform chirp impulse response given by

$$h(t) = \exp[j2\pi(f_o t + \tfrac{1}{2}kt^2)], \qquad -T/2 \le t \le T/2.$$

As in chapter three, $k = B/T$. Solving equation II.1 suggests that the transducer electrodes should be placed at positions such that

$$2(f_o t_n + \tfrac{1}{2}k\, t_n^{\,2}) = (n - \tfrac{1}{2}N_1) \quad ,$$

where $N_1 = 2f_oT+1$ is the total number of electrodes in the IDT (Gerard et al., 1973). N_1 is included to force $t_n = 0$ at the IDT centre. Solving for t_n gives the positions of the electrodes as

$$z_n = v_{eff}\,\frac{f_o}{k}\left\{-1 + [1 - \frac{k}{f_o^{\,2}}(\frac{N_1}{2} - n)]^{\frac{1}{2}}\right\} \qquad \text{II.4}$$

(Gerard, 1977).

Given the positioning rule for the IDT fingers it remains to determine the apodization (overlap) law for the transducer. In an ideal transversal filter the amplitude and phase weightings of the taps are determined solely by the sampled impulse response to be realised. If the filter taps are not completely independent it is necessary to pre-distort the filter design to account for the coupling.

In a SAW filter coupling can occur because of attenuation of the acoustic beam due to absorption by taps, imperfect electrode summation due to a non-zero load impedance terminating the IDT summing bus, or because of reflections within the device.

Gerard (1978) has applied a "crossed field" model to IDT design in an effort to account for coupling produced by the first two mechanisms. In the model each one-electrode section of the transducer is represented as a three-port network having one electric and two acoustic ports. By considering a cascade of these networks it is shown that the appropriate law for the electrode apodization is

$$w_n/W_o = (f_n/F)^{-3/2} a(t_n)/A \qquad \text{II.5}$$

where A and F are normalization constants. A is the maximum value of the chirp impulse response and F is chosen so that the maximum value of w_n equals W_o, the aperture of the acoustic beam. The synchronous frequency for the n*th* electrode is denoted by f_n. As expected, the overlap function is proportional to the required impulse response. However, the overlap length also depends on the local synchronous frequency and, since electrode spacing is not constant in a chirp device, the apodization must be adjusted accordingly.

The design equations II.4 and II.5 are sufficient to realise a practical filter if reflections within the IDT are small and if acoustic diffraction effects are negligible. In real filters these criteria are often met by combining a double electrode structure with a dummy electrode pattern (figure II-3). The double electrode arrangement maintains the required excitation gap periodicity but ensures that reflections from adjacent electrodes cancel since the total propagation phase shift of the reflected component is 180^o. Dummy electrodes do not affect the function of the IDT but prevent steering of the acoustic beam due to variation of the SAW velocity across the transducer aperture.

As a further design consideration it is worth noting that, for large time-bandwidth product devices, it is preferable to fabricate the DDL's with a down-chirp response and to negate one chirp slope in a mixer (a process referred to as spectral inversion) prior to

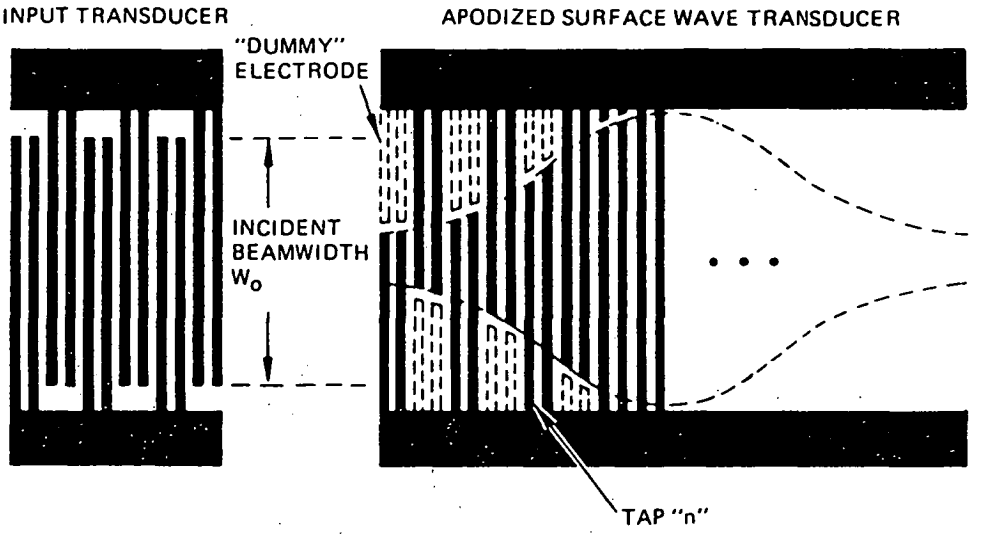


Fig.II-3: SAW filter incorporating a double electrode structure and a dummy electrode pattern. (After Gerard, 1978).

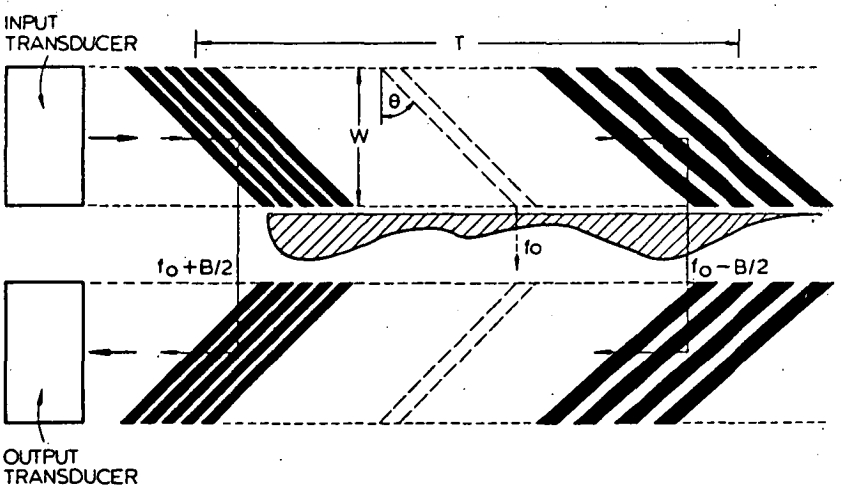


Fig.II-4: Basic RAC dispersive delay line. The shaded region between the two tracks is a deposited metal film for phase compensation. (After Ash, 1978).

compression. Ash (1978) shows that wideband down-chirp filters are less prone to spurious acoustic bulk wave responses than up-chirp devices and Schulz and Holland (1973) note that it is possible to achieve inherent temperature compensation in a down-chirp DDL (see also section 4.5).

Gerard (1977) shows that the input admittance of the electric port in the crossed field model is a function of frequency and that the largest practical transducer bandwidth achievable before phase and amplitude errors become excessive increases with increasing substrate piezo-electric coupling constant. Fractional bandwidths of 5% for quartz and 25% for lithium niobate ($LiNbO_3$) are representative upper limits for dispersive delay lines using a single chirp transducer.

When wider bandwidths are needed it is usual to employ two chirp transducers in a chirp to chirp arrangement. In this case the time signal to be synthesized at each transducer is

$$h_1(t) = F^{-1}\left\{[F(h(t))]^{\frac{1}{2}}\right\} ,$$

where F and F^{-1} denote forward and reverse Fourier transformation respectively and h(t) is the required impulse response (Gerard et al., 1973).

Detailed analyses of IDT behaviour (for example, Bristol, 1973 and Gerard, 1978) indicate that the design of SAW filters requires the designer to balance conflicting requirements such as insertion loss, bandwidth, phase error and device acoustic reflection level. In general, it is necessary to increase the insertion loss (either by mismatching to the IDT's or by using a substrate with a low piezo-electric coupling constant, such as quartz) to reduce phase error and reflection, including the triple transit reflection discussed in appendix one.

SAW DDL's exhibit an insertion loss which is related directly to time-bandwidth product since the need to minimize phase error increases with TB. In addition, as the complexity of the metallization increases reflection suppression becomes more difficult and the techniques adopted often lead to increased insertion loss.

Further, most large TB devices use a quartz substrate since the temperature dependence of the acoustic delay is then less than with say, $LiNbO_3$. Unfortunately, the piezo-electric coupling constant is also reduced and the filter insertion loss again increases. Insertion losses of about 30 dB and 55 dB are typical for DDL's with TB = 50 and TB = 1000 respectively. The new low-loss polyphase IDT structure referred to in appendix one has not yet been applied in commercial DDL's, principally because of the increased metallization complexity.

As the result of research aimed at reducing the insertion loss and increasing the response accuracy of large TB DDL's, a second SAW realization is now available. This device, called a reflective array compressor (RAC), is described in detail by Williamson (1973) and is illustrated schematically in figure II-4. Conventional wideband (few fingers) IDT's are used to generate and receive Rayleigh waves. Dispersion occurs because of the spatially selective reflection from an acoustic grating formed by slots in the substrate surface. The path of a given frequency component is U-shaped since the input propagates to a region where the slot spacing is synchronous, is reflected to the other half of the device and again reflected towards the receiving IDT.

Ash (1978) identifies several advantages of the RAC device. These are:

(a) the substrate length needed for a given total dispersion is halved by comparison with an in-line IDT;

(b) spurious responses are suppressed since responses due to bulk acoustic waves or unwanted reflections are unlikely to follow the complicated path necessary for the waves to arrive at the receiver IDT with wavefronts parallel to the fingers;

(c) interaction between acoustic and electric sections of the filter is minimised since the reflecting slots are not connected to an electrical circuit;

(d) phase compensation of the DDL is possible since each frequency component follows a unique path. The compensation is provided by a longitudinal thin-film "phase plate" which modifies the SAW velocity between the "U" sections;

(e) amplitude weighting can be provided by varying the depth of the slots in the grating, an easy process when ion beam fabrication methods are used.

RAC filters with TB~10 000 are available compared with an upper limit of TB~1000 for in-line DDL's. The electrical performance equals or exceeds that of conventional SAW devices and developments such as "in-line" and "dotty" RAC's (Plessey Research, 1981) offer still better results with the possibility of reduced price relative to standard RAC's (due to simpler manufacturing methods).

References

Ash, E.A.: Fundamentals of signal processing devices - in "Acoustic surface waves" (A. Oliner ed.), Ch.4, Springer-Verlag, 1978.

Bristol, T.W.: Analysis and design of surface acoustic wave transducers. IEE Conf. Pub.109, 1973, pp.115-129.

Farnell, G.W.: Elastic surface waves - in "Surface wave filters" (H. Mathews ed.), Ch.1, Wiley, 1977.

Gerard, H.M.: Surface wave interdigital chirp filters - in "Surface wave filters" (H. Mathews ed.). Ch.8, Wiley, 1977.

Gerard, H.M.: Principles of surface wave filter design - in "Acoustic surface waves" (A. Oliner ed.), Ch.3, Springer-Verlag, 1978.

Gerard, H.M., Smith, W.R., Jones, W.R. and Harrington, J.B.: The design and applications of highly dispersive acoustic surface wave filters. IEEE Trans. Son. Ultrason., Vol.SU-20, No.2, April 1973, pp.94-104.

Plessy Research (Caswell) Ltd.: Background Notes: surface acoustic wave symp., February 1981.

Schulz, M.B. and Holland, M.G.: Materials for surface acoustic wave components. IEE Conf. Pub.109, 1973, pp.1-10.

Williamson, R.C.: Large time-bandwidth product devices achieved through the use of surface acoustic wave reflection gratings. IEE Conf. Pub.109, 1973, pp.115-129.

APPENDIX III - GENERATION OF RANDOM NUMBERS WITH A GIVEN PROBABILITY DISTRIBUTION

Most computer systems give programmers access to a random number generator. One of the most common and satisfactory methods of generation is to use a multiplicative congruence generator (Hamming, 1973) to produce uniformly distributed numbers. By convention the distribution is normalised such that $0 \leq x_i < 1$, where x_i is the random number. Random numbers used as the basis for plots in this appendix and as data for the simulations described in chapter five were generated using the RT-11 FORTRAN IV function RAN(I,J). The RAN(I,J) function calls the library subroutine RANDU which implements a multiplicative congruence generator.

Hamming describes a method of generating random numbers distributed according to other probability density functions. The method allows a given number, x_i, from the standard uniform distribution to be transformed, yielding a second number, y_i, from the desired distribution. A more general description of random number transformation is given by Zelen and Severo (1965).

To determine the transformation formula the cumulative probability distribution, F(y), of the wanted p.d.f. is equated to the cumulative distribution, x, of the uniform distribution. That is,

$$\int_0^x 1.dx = x = \int_0^y p(y)dy = F(y). \qquad \text{III.1}$$

Suppose for example that exponentially distributed random numbers are needed. In this case the p.d.f. is

$$p(z) = \frac{1}{\sigma^2} \exp(-z/\sigma^2)$$

and the cumulative density function is

$$F(z) = \frac{1}{\sigma^2} \int_0^z \exp(-z/\sigma^2)\, dz = 1 - \exp(-z/\sigma^2).$$

Equating cumulative density functions gives

$$x = 1 - \exp(-z/\sigma^2)$$

and $$z = -\sigma^2 \ln(1-x) \ .$$

Thus,

$$z_i = -\sigma^2 \ln(1-x_i).$$

Since $0 \leq x_i < 1$, an equally effective generating formula is

$$z_i = -\sigma^2 \ln x_i \ . \qquad \text{III.2}$$

A plot of the class distribution of 10 000 numbers ($\sigma = 1$) obtained using this formula is shown in figure III-1. The theoretical distribution is also shown.

As a second example, suppose that Rayleigh distributed numbers are required. In this case

$$p(V) = \frac{V}{\sigma^2} \exp\left(-\frac{V^2}{2\sigma^2}\right)$$

and

$$F(V) = \frac{1}{\sigma^2} \int_0^V V \exp\left(-\frac{V^2}{2\sigma^2}\right) dV.$$

By direct differentiation,

$$\int V \exp(-V^2/a)\ dV = -\frac{a}{2}\exp(-V^2/a) + \text{constant}.$$

Hence,

$$F(V) = 1 - \exp\left(-\frac{V^2}{2\sigma^2}\right) \ .$$

Equating cumulative density functions gives

$$x = 1 - \exp\left(-\frac{V^2}{2\sigma^2}\right)$$

so

$$V = \sqrt{-2\sigma^2 \ln(1-x)}$$

or, for $0 \leq x_i < 1$,

$$V_i = \sqrt{2}\ \sigma\ \sqrt{-\ln x_i} \ . \qquad \text{III.3}$$

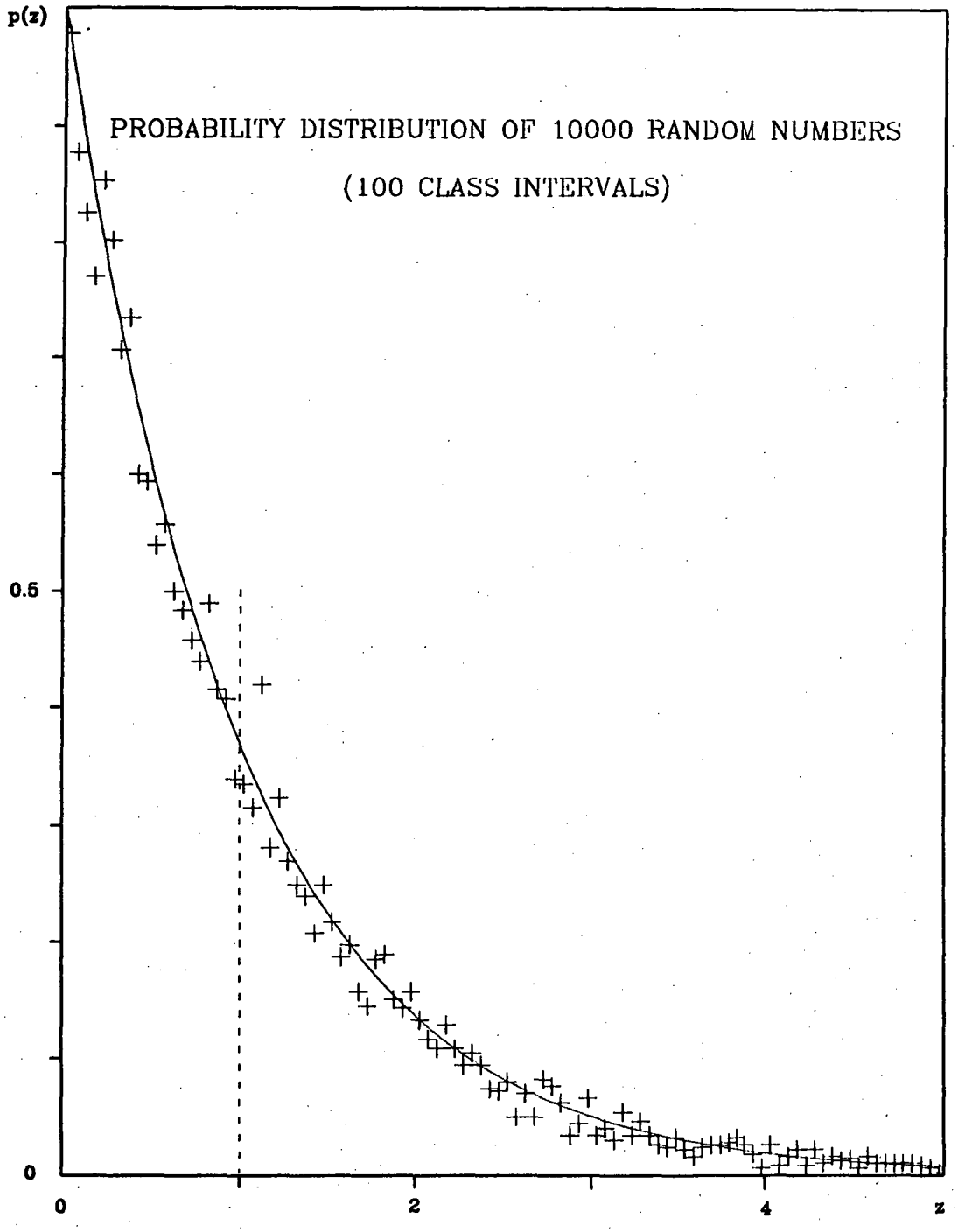


Fig.III-1: Distribution of N = 10 000 random numbers generated using equation III.2(σ = 1). The distribution is plotted as a normalized histogram having 100 class intervals of width a = 0.05. The solid line shows the theoretical exponential p.d.f.,p(z), and the dashed line indicates the mean of this distribution. The mean and standard deviation of the generated numbers are within 1 part in 10^3 of the theoretical values. Using a Gaussian approximation to the binomially distributed number of random variables falling in a given interval, it can be shown that the expected 1σ local scatter about the curve is $\sim \sqrt{\frac{p(z_i)[1-p(z_i)]}{Na}}$, where $p(z_i)$ is p(z) at the i*th* interval mid-point. For example, the 1σ scatter for $z \sim 1$ should be ~ 0.02 - close to the observed spread. Similar agreement holds in other regions.

Figure III-2 is a plot for $\sigma = 1$ of the class distribution of 10 000 numbers derived using this transformation. Again, the fit to the expected distribution is good.

In some cases random numbers with a given mean are required. For the exponential distribution $\bar{z} = \sigma^2$, so equation III.2 can be applied directly to give

$$z_i = -\bar{z} \ln x_i \ .$$

For the Rayleigh distribution

$$\bar{V} = \sqrt{\frac{\pi}{2}}\,\sigma$$

so

$$\sigma = \bar{V}\sqrt{\frac{2}{\pi}}$$

and

$$V_i = \frac{2}{\sqrt{\pi}}\,\bar{V}\sqrt{-\ln x_i} \ .$$

In both cases the random numbers can be generated with unit mean and scaled according to requirements.

Hamming makes the point that it is easier to generate useful, nearly independent, random numbers (such as those required in section 5.3) than to generate trains of numbers with intentional internal correlations. As far as electrical noise simulation is concerned, plots such as those of figures III-1 and III-2 give confidence in the generation process.

References

Hamming, R.W.: "Numerical methods for scientists and engineers". Ch.8, McGraw-Hill, 1973.

Zelen, M. and Severo, N.C.: Probability functions-in "Handbook of mathematical functions" (M. Abromovitz and I. Stegun ed.), Ch.26, Dover, 1965.

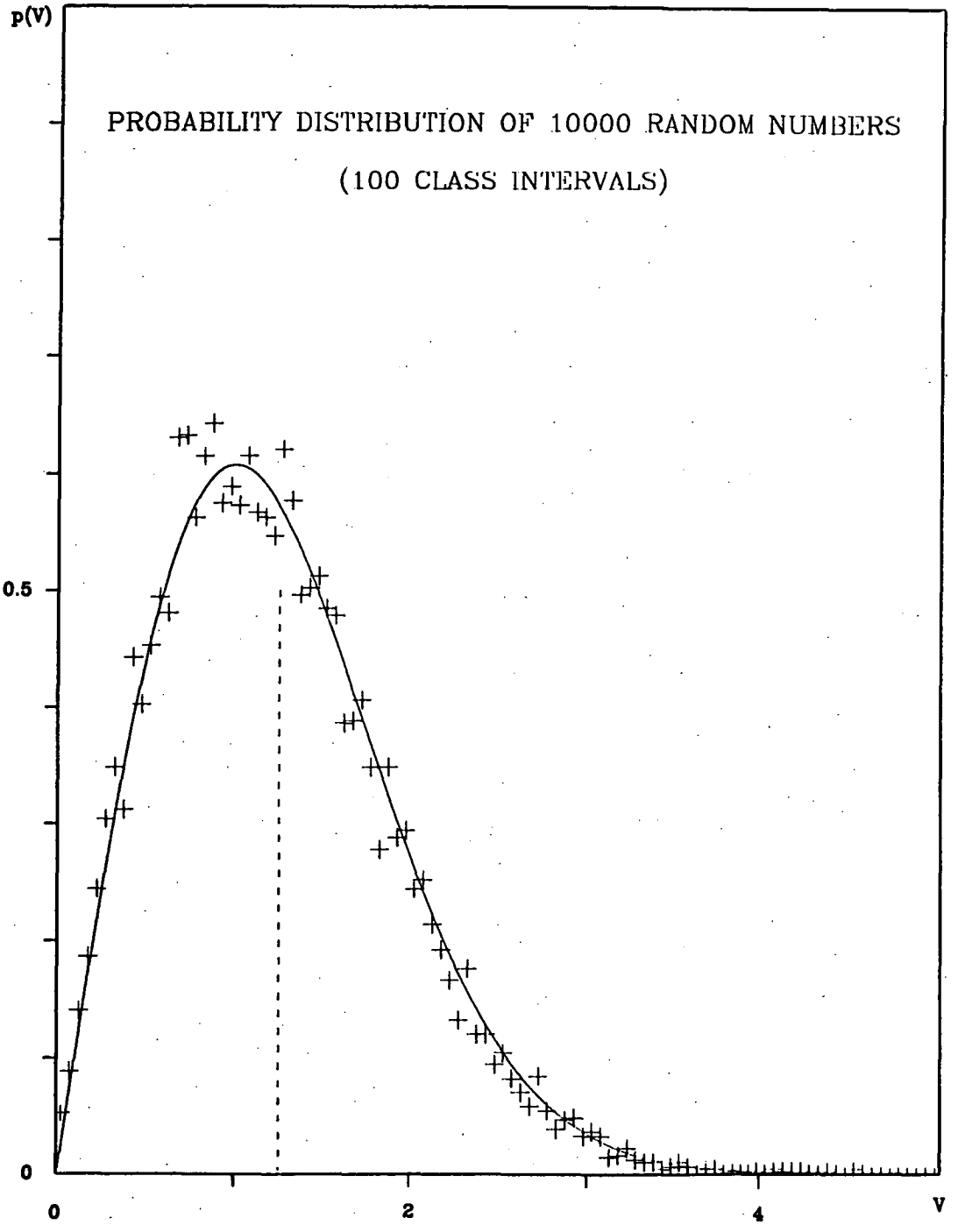


Fig.III-2: Distribution of 10 000 random numbers generated using equation III.3 (σ = 1). The solid curve shows the theoretical Rayleigh p.d.f.,p(V); the dashed line indicates the mean of p(V). Comments concerning the style of data presentation, accuracy of the generated distribution and scatter of points in figure III-1 also apply to the above plot.

APPENDIX IV MARK 1 DVI CONTROLS

The control panel of the enclosure containing the four "mark 1" DVI's is shown in figure 8-6. The following summary describes briefly the function of each control and connector. An asterisk denotes a connector or control associated with the "pulsar" mode of operation discussed in section 10.2 and which has no effect on the "normal" operational mode considered in chapter eight. The panel is designed so that for pulsar observations the control switches are all down and all the LEDS are extinguished. All digital input and output signals are TTL compatible.

DVI CONTROLS AND CONNECTORS

(1) "Set N" thumbwheel switches - Selects the number of integrations to be performed by the DVI's.

(2) "Mode" toggle switch - Selects "normal" or "pulsar" mode.

(3) "Trigger source" toggle switch - Selects "internal" or "external" trigger.

(4) "Display clock" toggle switch - Selects the source of the buffer RAM READ address clock. "Local" enables an internal oscillator, "line" allows a data acquisition computer access to the four buffers.

(5) "Output gain" rotary switch (4) - Selects which eight of 16 bits in the output buffer RAM are presented to the Y-axis D/A converter when a CRT display of integrated spectral power versus frequency cell is in use.

(6) " Output offset" toggle switch (4) - Allows the Y-axis D/A converter output to be offset. Selecting the "off" position enables a rapid assessment of absolute levels to be made.

(7) " Output offset" ten-turn potentiometer (4) - Sets the Y-axis D/A converter d.c. offset.

(8) "Video input" BNC (4) - A/D converter input. Signals in the range -1 to 0 V are assumed.

(9) "Video output" BNC (4) - Y-axis output for integrated spectrum display. D/A converter range is ± 5 V.

(10) "Horizontal output" BNC - X (or frequency) axis output for integrated spectrum display. D/A converter range is ± 5 V.

(11) "External trigger input" BNC - When "external" trigger mode is selected the DVI control module is triggered on the first high-going DVI master clock edge following a high-going transition at this input.

(12) "Trigger output" BNC - SAW spectrometer trigger pulse output. The pulse width is half the master clock period and the signal is buffered by 50 ohm line drivers.

*(13) "Go" BNC - An output signal sent directly from the computer to the pulsar timing unit (PTU). The signal has no direct effect on the DVI's.

*(14) "Pulsar out" BNC - A replica of the DVI control module "new data available" pulse. The signal allows the PTU to count the number of data transfers from the DVI's to a data acquisition computer since the last PTU "delayed frame" signal.

*(15) "Pulsar in" BNC - An input signal which allows synchronization of the integration process with the PTU "delayed frame" signal. Synchronization is achieved by forcing the DVI control module to be triggered an integral number of times during the pulsar period.

(16) "Local display clock" LED - Warns that internal video display mode has been selected and that computer access to the data buffers of the four DVI's is impossible.

(17) "N too small" LED - Indicates that a data acquisition computer received a "new data available" interrupt before complete transfer of integrated spectral data was accomplished. A larger number of integrations must be set for satisfactory transfer to take place.

(18) "External trigger" LED - Warns that external trigger mode has been selected.

APPENDIX V A LARGE-SCALE SAW RADIOMETER

In this appendix the specifications of a second SAW radiometer now under development are given. Although the author gave extensive advice during the design phase, the dual-channel SAW processor and its four associated "mark 2" DVI's were constructed by G.R.A. Ellis and G.A. Gowland. At present, the scientific program envisaged for the instrument does not require correlation of the two spectrum analyser outputs. However, good results were obtained when the VHF correlator (section 6.7) was used as a temporary post-transformation processor. Characteristics of the basic SAW analyser are listed in table V-1.

Table V-1 SAW Spectrum Analyser Specifications

Analyser type:	(C)-M-C, identical DDL's used in convolver and LO chirp generator sections.
DDL type:	Andersen Laboratories DS1310
- chirp sign:	Negative (down-chirp).
- bandwidth (B_1):	19 MHz.
- chirp duration (T_1):	34 μs.
- time-bandwidth product:	646.
- centre frequency:	91 MHz.
- weighting:	None.
Analyser centre frequency:	250 MHz.
Duty cycle:	50% (initially).
LO chirp generator:	Active type based on two DS1310 DDL's (described in section 6.5 design example).

Characteristics of the "mark 2" DVI are listed in table V-2.

Table V-2 Mark 2 DVI Specifications

Sample frequency:	20 MHz.
Quantization resolution:	6 bits (TRW TDC1014J A/D converter).
Number of frequency cells:	1024.
Buffer readout options:	Local (CRT) at present but compatible with DMA interface described in section 10.3.

The TDC1014J A/D converter has similar short-comings to the TDC1021J four-bit device (section 8.9). However, the newer TDC1029 converter (chapter 12) should give improved performance. The "mark 2" DVI is functionally identical to the "mark 1" unit described in chapter eight. Internally however, a fast serial-to-parallel converter replaces the ALU shift register circulating memory. The basic architecture is shown in figure V-1. In this design the data rate in the individual arithmetic stages is only 1/8*th* of the input data rate. Thus, "LS" TTL logic elements and ordinary MOS memories can be used in each stage.

Although more attractive for large-scale DVI's, this architecture requires a considerably more sophisticated control system than the shift register approach. In economic terms it offers no advantage until the number of frequency cells exceeds approximately 256.

Figure V-2 shows some preliminary data obtained using one of the new spectrum analysers and its associated DVI. The 1024 cell output of the DVI was recorded on moving photographic film, with increasing intensity producing a darker image in figure V-2. The useful frequency range of the receiving system was 8-35 MHz. In figure V-2, interference from short-wave broadcasting stations is severe at lower frequencies but Jovian "L" bursts are very clear in the region around 20 MHz. The bursts were recorded using a 16 element crossed-dipole array at the Llanherne observatory. Left and right-hand circular polarizations were formed using a broadband quadrature hybrid. The spectrum analyser input was switched between the two modes at intervals of several milli-seconds. Diode linear detectors were used at the spectrum analyser output. The frequency-time slope of the signals is of the order of hundreds of kilohertz per second and is typical of this class of burst emission.

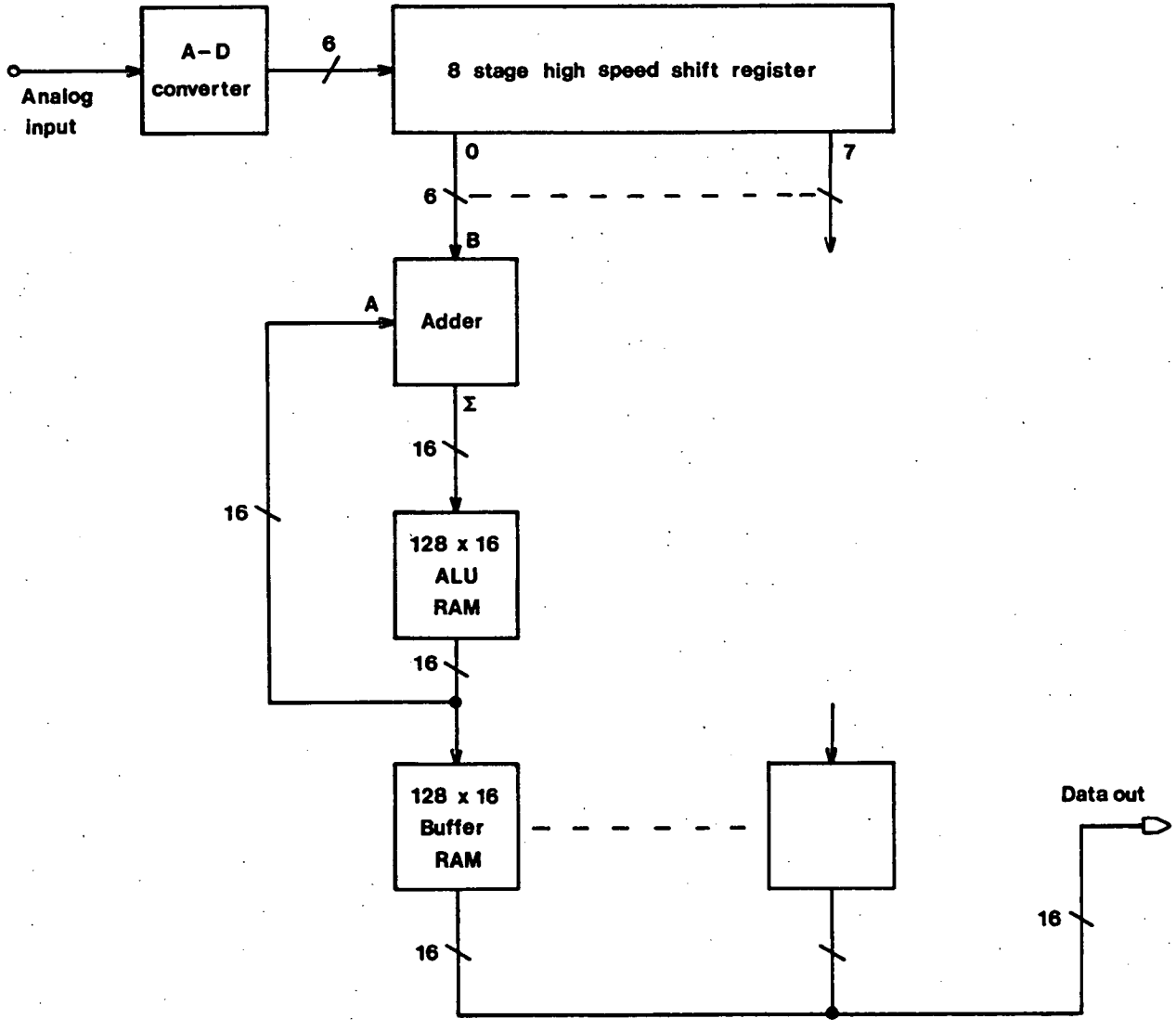


Fig.V-1: Architecture of the "mark 2" DVI. Data from the A/D converter are clocked serially into the high speed shift register. When the register is full, its contents are transferred sideways to the eight parallel arithmetic units.

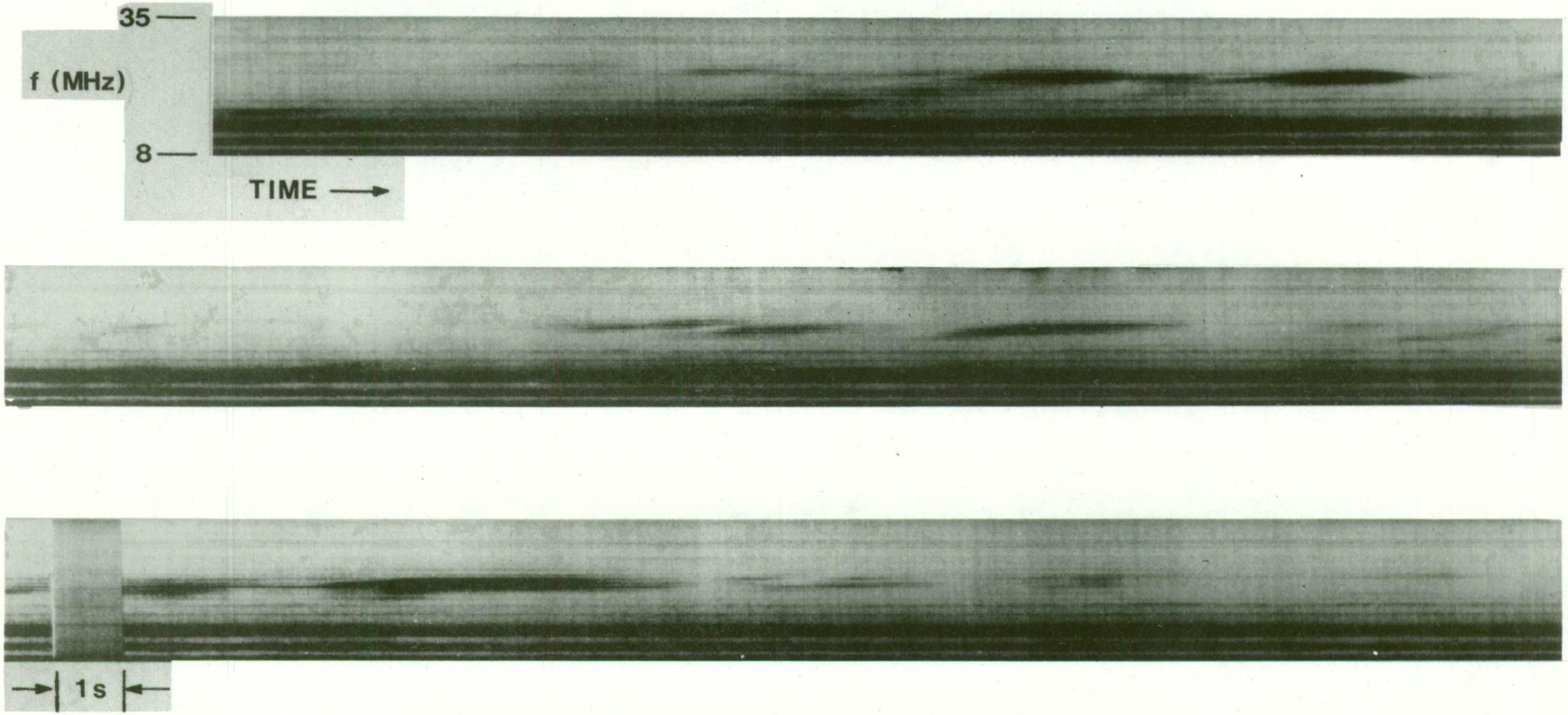


Fig.V-2: Jovian "L" bursts recorded at Llanherne on 29 July 1984 using the SAW radiometer described in the text. (Data courtesy of Professor G.R.A. Ellis).

APPENDIX VI LIST OF PUBLICATIONS

The following publications have been produced wholly or substantially by the author during the course of this project. The list is complete to 30 January 1985.

1. Refereed Publications

Hall, P.J.: A multiplier/detector for SAW spectrum analysers. Electron. Lett., Vol.18, No.16, Aug. 1982, pp.699-700.

*Hall, P.J., Hamilton, P.A. and McCulloch, P.M.: A correlation Stokes polarimeter using a dual channel SAW Fourier analyser. JEEE (Aust.), Vol.2, No.3, Sept. 1982, pp.150-155.

Hall, P.J., Hamilton, P.A. and McCulloch, P.M.: A radio polarimeter spectrometer, Proc. Astron. Soc. Aust., Vol.5, No.2, 1983, pp.260-262.

* - Paper awarded the Norman W.V. Hayes Memorial Medal of the Institution of Radio and Electronic Engineers (Aust.) for the most meritorious original research paper published in the IREE "Proceedings" during 1982.

2. Conference Paper

Hall, P.J., Hamilton, P.A., Lyne, A.G. and McCulloch, P.M.: SAW spectrum analysis in radioastronomy. Proc. IEE Coll. "SAW Devices and Systems", London, 1982.

3. Other

Hall, P.J.: Surface acoustic wave devices - fundamentals and applications. Electronics Today International, Nov. 1983 and Jan. 1984, pp.105-109 and 118-120. (Reprinted as appendix one in this thesis).

APPENDIX VII A NOTE ADDED IN PROOF RELATING TO THE EFFECT OF SPECTRUM ANALYSER SAMPLING LOSS ON RADIOMETER EFFICIENCY

Pryor (1973) considers the problem of sampling and averaging the un-smoothed output of the square-law detector in a narrowband channel (figure 5-1). He presents the results of a study undertaken by him within the U.S. Naval Ordnance Laboratory and described in an internal report of that organization. The study assumes ideal sampling and does not address the problem of quantization using finite resolution A/D converters. However, Pryor's work quantifies the reduction in SNR which occurs due to the use of a finite sample frequency at a point between the detector and integrator. The detector output would normally be sampled at a frequency at least equal to Δf, the actual sample frequency (or frequency of generation of spectral estimates) relative to Δf being referred to as the "redundancy" of the system. Strictly, Δf is an equivalent bandwidth and is equal to the -3 dB bandwidth only for certain IF filter responses. Sampling at Δf results in no loss in SNR at the integrator output only if the IF passband is rectangular. For a SAW analyser $\Delta f = 1/T_1$.

Pryor's results for un-weighted and Hanning weighted correlation or FFT spectrum analysers can be applied directly to equivalent SAW systems. If the resolution cell shape in the prototype analysers is assumed to be similar to that produced by a Hanning weighting function, the expected radiometer efficiency can be calculated. The assumption of passband similarity is reasonable: the Hanning function produces a -3 dB pulse broadening factor of 1.6 compared with the pre-detection factor of 1.5 in the actual analysers.

The redundancy of the prototype processor is 0.6, assuming 4 μs spectrometer and integrator triggering (60% duty cycle). The sampling loss thus incurred is about 2.7 dB, resulting in a predicted radiometer efficiency of 54%. Using the results of section 9.2, the actual loss is (17.9 - 15.4) or 2.5 dB, giving an efficiency of 56% if no allowance for non-continuous operation is made. If a quantizing loss of 0.5 dB is assumed, the projected maximum efficiency of the prototype radiometer is 63%. The measured performance is therefore fairly close to that expected on the basis of Pryor's study. At 100%

duty cycle the predicted sampling loss is about 1.6 dB, giving a radiometer efficiency of 70%. The equivalent figures for an un-weighted analyser operating at 100% duty cycle are 0.9 dB and 82%. The figures for a continuously operating SAW radiometer are identical to those applying to, for example, an FFT radiometer (spectrum analyser and integrator) utilizing a single spectrum analyser. Like a SAW system, such a radiometer has unity redundancy because the integrator receives only one spectrum per block of input data.

Reference

Pryor, C.N.: Minimum detectable signal for spectrum analyser systems. Proc. NATO Adv. Stud. Inst. on Sig. Proc. (Underwater Acoust.), Academic Press, 1973.